# DOCUMENT RESUME

ED 335 418 TM 017 119

| | |
|---|---|
| AUTHOR | Dowd, Kathryn L.; And Others |
| TITLE | The NELS:88 Second Follow-Up Field Test Report. Volume 1. |
| INSTITUTION | National Opinion Research Center, Chicago, Ill. |
| SPONS AGENCY | National Center for Education Statistics (ED), Washington, DC. |
| PUB DATE | Jun 91 |
| CONTRACT | NCES-RS90005001 |
| NOTE | 344p.; For Volume 2, which contains field test questionnaires and data, see TM 017 120. |
| PUB TYPE | Reports - Descriptive (141) |
| EDRS PRICE | MF01/PC14 Plus Postage. |
| DESCRIPTORS | Cognitive Tests; Cohort Analysis; Dropouts; *Field Tests; *Followup Studies; High Schools; High School Students; Junior High Schools; *National Surveys; Parents; Principals; *Questionnaires; Research Methodology; *Test Construction |
| IDENTIFIERS | Free Response Test Items; *National Education Longitudinal Study 1988 |

ABSTRACT

The Second Follow-Up Field Test (SFUFT) of the National Education Longitudinal Study of 1988 (NELS:88) was conducted in 1990 and 1991 by the National Opinion Research Center and the Educational Testing Service. The SFUFT evaluated free response items (FRIs) for possible inclusion in the Second Follow-Up Main Study (SFUMS) and tested survey instruments, procedures, and forms. This report focuses on a methodological examination of the SFUFT data and the questionnaires developed. The SFUFT included four major survey components: students, dropouts, parents, and school administrators. Four questionnaires and five cognitive tests were developed, including the: Student Questionnaire, Not Currently In School (Dropout) Questionnaire, Parent Questionnaire, School Administrator Questionnaire (SAQ), and Cognitive Test Battery. Most of the students surveyed were in the last half of their senior year in high school. The major purpose of the parent component field test was to gather sufficient observations for evaluating the performance of questionnaire items. Data were obtained from: 2,254 students in 94 schools (overall response rate of 70.5%); 108 dropouts; 506 parents (response rate of 73.5%); and 65 schools/principals. The results show that the goals set for this pretest effort were attained, the FRIs were successfully administered, and the FRIs measure the same/similar domains as do multiple-choice questions. Recommendations for implementation of the SFUMS are provided. Numerous tables and figures are included. Documentation of 1990 instrument development meetings, documentation supporting questionnaire experiment analyses and cognitive test analyses, and a summary of cognitive interviews are appended. (RLC)

***********************************************************************
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
***********************************************************************

ED335418

4521/4525

Volume 1

# THE NELS:88 SECOND FOLLOW-UP FIELD TEST REPORT

**KATHRYN L. DOWD**

**STEVEN J. INGELS**

**JUDITH M. POLLACK**

**KENNETH A. RASINSKI**

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Steven Ingels

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

in collaboration with: Deborah Alexander, Eric Camburn, Donald A. Rock, Timothy Ryan, John R. Taylor, Lisa Thalji and Roger Tourangeau.

Prepared by the National Opinion Research Center (NORC) at The University of Chicago. Prepared for the Longitudinal and Household Studies Branch of the National Center for Education Statistics, U.S. Department of Education.

Steven Ingels, Project Director; Peggy Quinn NCES Project Officer

CHICAGO, JUNE 1991

The National Education Longitudinal Study of 1988 (NELS:88) Second Follow-Up is conducted by NORC for the National Center for Education Statistics under contract No. RS90005001.

TM017119

BEST COPY AVAILABLE

## NELS:88 Second Follow-Up Field Test Report

## Table of Contents

Author List

Executive Summary

Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.0 Field Test Preparations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

- 1.1 Development of Field Test Instruments . . . . . . . . . . . . . . . . . . . . 9
- 1.2 Field Staff Recruiting and Training . . . . . . . . . . . . . . . . . . . . . . . 14
- 1.3 Sample Design and Selection . . . . . . . . . . . . . . . . . . . . . . . . . . 16
- 1.4 Student Tracing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
- 1.5 Securing State, District, and School Cooperation . . . . . . . . . . . . . 23
- 1.6 Obtaining Parental Permission . . . . . . . . . . . . . . . . . . . . . . . . . . 26
- 1.7 Identification of Dropouts and Early Graduates . . . . . . . . . . . . . . 27

2.0 Student Survey Procedures . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29

- 2.1 Objectives, Content, and Experiments . . . . . . . . . . . . . . . . . . . . 29
- 2.2 Field Test Data Collection Procedures . . . . . . . . . . . . . . . . . . . . 31
- 2.3 Data Preparation Procedures . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
- 2.4 Recommendations for Changes in Student Data Collection Procedures . . . . . . . . . . . . . . . . . . . . . 38

3.0 Analysis of Student Questionnaire and Test Data . . . . . . . . . . . . . . . . . . 41

- 3.1 Student Response Rates . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41
- 3.2 Item Nonresponse Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
- 3.3 Logical Consistency of Response to Filter and Dependent Questions . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60
- 3.4 Results of Student Questionnaire Experiments . . . . . . . . . . . . . . . 65
- 3.5 Item Nonresponse By Item Postion . . . . . . . . . . . . . . . . . . . . . . 73
- 3.6 Scale Reliabilities . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 74

Table of Contents - continued

3.7 Student Data Quality .................................... 78
3.8 Response to OMB Queries - Student Questionnaire ............ 80
3.9 Recommendations for Student Questionnaire ................ 87
3.10 Preliminary Analysis of the Cognitive Test Battery ........... 91

4.0 Dropout Survey Procedures ................................ 115

4.1 Objectives and Content ................................ 119
4.2 Defining Dropout Sample Members ........................ 121
4.3 Group Administration Data Collection Procedures ............. 121
4.4 Data Preparation ...................................... 122
4.5 Recommendations for Changes in Dropout
Data Collection Procedures ............................. 123

5.0 Dropout Survey Data Analysis ............................. 127

5.1 Dropout Questionnaire Response Rates .................... 127
5.2 Analysis of the Dropout Questionnaire ..................... 127
5.3 Response to OMB Queries - Dropout Questionnaire ............ 139
5.4 Recommended Changes to the Dropout Questionnaire .......... 140

6.0 Parent Survey Procedures ................................ 143

6.1 Objectives, Content, and Experiments ..................... 143
6.2 Sample Selection for the Parent Survey ................... 145
6.3 Data Collection ....................................... 146
6.4 Data Preparation ...................................... 148

7.0 Results of Parent Data Analysis ............................ 151

7.1 Parent Response Rates ................................ 151
7.2 Parent Item Nonresponse .............................. 151
7.3 Results of Parent Questionnaire Experiments ................ 162

Table of Contents - Continued

7.4 Logical Consistency of Responses to Filter and Dependent Questions . . . . . . . . . . . . . . . . . . . . . . . . . . . . 168

7.5 Parent Data Quality Assessment . . . . . . . . . . . . . . . . . . . . . . . . . . 170

7.6 Results of Parent Survey Mail Experiments . . . . . . . . . . . . . . . . . 171

7.7 Response to OMB Queries - Parent Questionnaire . . . . . . . . . . . . . 172

7.8 Recommendation for Parent Survey . . . . . . . . . . . . . . . . . . . . . . . 173

8.0 School Administrator Survey Procedures . . . . . . . . . . . . . . . . . . . . . . . . . . 177

8.1 School Administration Questionnaire: Objectives and Content . . . . 177

8.2 Respondent Selection and Data Collection . . . . . . . . . . . . . . . . . . 178

8.3 Data Preparation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 179

9.0 Results of School Survey Data Analysis . . . . . . . . . . . . . . . . . . . . . . . . . 181

9.1 School Administrator Questionnaire Response Rate . . . . . . . . . . . . 181

9.2 Item Nonresponse Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 181

9.3 Analysis of Response to Filter and Dependent Questions . . . . . . . . 188

9.4 Responses to OMB Queries - School Administrator Questionnare . . 189

9.5 Recommendations for Changes for the School Survey . . . . . . . . . . 192

10.0 Preliminary Assessment of Course Offerings and Enrollment Data Collection Procedures . . . . . . . . . . . . . . . . . . . . . . 195

10.1 Course Offerings Data Collection Procedures . . . . . . . . . . . . . . . . 195

10.2 Enrollment Data Collection Procedures . . . . . . . . . . . . . . . . . . . 196

10.3 Preliminary Recommendations for Course Offerings and Enrollment . . . . . . . . . . . . . . . . . . . 197

11.0 Summary of Recommendations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 199

11.1 Summary of Recommendations for Changes in Procedures . . . . . . 199

11.2 Recommendations for Changes in the Instruments . . . . . . . . . . . . 200

11.3 Recommendations for the Teacher Survey . . . . . . . . . . . . . . . . . . 202

Table of Contents - Continued

11.4 Recommendations for the Early Graduate Supplement .......... 207

11.5 Further Plans in the NELS:88 2FU for Addressing the Problem of Excluded Students .............................. 208

11.6 Issues Related to the School Effects Supplement ............. 210

REFERENCES ............................................... 213

APPENDIX A

Documentation of 1990 Instrument Development Meetings

APPENDIX B

Documentation in Support of Questionnaire Experiment Analyses

APPENDIX C

Supporting Documentation for Cognitive Test Analyses

APPENDIX D

Summary of Cognitive Interviews

APPENDIX E

Field Test Questionnaires and Data

# LIST OF AUTHORS

## Volume 1

| | | |
|---|---|---|
| Executive Summary | | Steven Ingels<br>Katy Dowd |
| Introduction | | Katy Dowd<br>Steven Ingels |
| Chapter 1 | Field Test Preparations | Eric Camburn<br>John Taylor<br>Steven Ingels<br>Katy Dowd |
| Chapter 2 | Student Survey Procedures | Katy Dowd<br>John Taylor<br>Steven Ingels |
| Chapter 3 | Results of Student Data Analysis | Ken Rasinski<br>Eric Camburn<br>John Taylor<br>Roger Tourangeau |
| | Preliminary Analysis of Cognitive Test Battery | Judy Pollack<br>Don Rock |
| Chapter 4 | Dropout Survey Procedures | John Taylor<br>Katy Dowd<br>Steven Ingels<br>Eric Camburn |
| Chapter 5 | Results of Dropout Data Analysis | John Taylor<br>Katy Dowd<br>Eric Camburn |
| Chapter 6 | Parent Survey Procedures | John Taylor<br>Ken Rasinski<br>Steven Ingels |
| Chapter 7 | Results of Parent Data Analysis | John Taylor<br>Eric Camburn<br>Tim Ryan<br>Ken Rasinski<br>Roger Tourangeau<br>Katy Dowd |

| | | |
|---|---|---|
| Chapter 8 | School Survey Procedures | John Taylor<br>Steven Ingels |
| Chapter 9 | Results of School Data Analysis | Eric Camburn<br>John Taylor<br>Katy Dowd |
| Chapter 10 | Preliminary Assessment of Course Offerings and Enrollment Data Collection Procedures | Debbie Alexander<br>Katy Dowd |
| Chapter 11 | Summary of Recommendations | Katy Dowd<br>Steven Ingels<br>Lisa Thalji |
| Appendix A | Documentation of 1990 Instrument Development Meetings | Leslie Scott<br>Steven Ingels<br>Anne Hafner |
| Appendix B | Documentation in Support of Student Questionnaire Experiment Analyses | Ken Rasinski |
| Appendix C | Supporting Documentation for Cognitive Test Analyses | Judy Pollack<br>Don Rock |
| Appendix D | Summary of Cognitive Interviews | Tim Ryan<br>Craig Hill<br>Gayle Hutchinson<br>David Mingay<br>Roger Tourangeau |

Volume 2

| | | |
|---|---|---|
| Appendix E | Field Test Questionnaires and Data<br>Student Questionnaire<br>Dropout Questionnaire<br>Parent Questionnaire<br>School Administrator Questionnaire<br>Proposed Early Graduate Supplemental Questions | Holly Chandler<br>Steven Ingels |

## EXECUTIVE SUMMARY OF THE FIELD TEST REPORT

The Field Test of the National Education Longitudinal Study of 1988 (NELS:88) Second Follow-Up was conducted in 1990 and 1991 by NORC and the Educational Testing Service (ETS). The field test cohort (located in California, Florida, Illinois, New York, and Texas) drawn in the Base Year and used in the First Follow-Up was surveyed, and an additional 2,500 high school seniors augmented the sample to produce a higher number of cognitive test observations needed for the assessment of free response items. The purpose of the 2FU Field Test was to evaluate free response items for possible inclusion in the main study, and to test survey instruments, procedures, and forms.

Many of the standard features of NELS:88, including student tracing and most data collection methodologies, have been thoroughly tested in the Base Year and First Follow-Up field tests and main studies. To the extent that this was so, the emphasis of the Second Follow-Up Field Test was often deliberately altered, so that new issues could be explored. Thus, for example, comparatively little attention was given to retesting methodologies for locating dropouts and surveying them in one-on-one administrations, since the First Follow-Up methodologies for doing so are available to the study and have proved sound. On the other hand, group sessions for dropouts have not been conducted in prior rounds of NELS:88, and it was felt that field test resources should be concentrated on assessing the group methodology.

Another example of such a change in emphasis may be seen in the cognitive tests. A major issue to be explored in the field test was the potential contribution of free response items to the NELS:88 cognitive test battery. Because of the reconfiguration of the testing design -- a much larger number of observations was required than could be supplied by the longitudinal field test cohort, and longitudinal data were not required as such -- field test sample selection and data collection were shaped accordingly. The longitudinal sample clusters were augmented to provide the needed number of test observations. In many respects, this redesign changed the nature of the field test sample, introducing a lack of strict comparability to the main study sample. In the main study, the 1500 NELS:88 clusters will average around a dozen students per school, with several hundred in the range of one to five students. In the field test, in-school sample sizes were generally triple the average number of students anticipated in the main study and dominated by augmented students with no prior experience of NELS:88. Nonetheless, procedures that are sensitive to variation in sample size (or membership in a longitudinal cohort) have been developed and thoroughly tried in earlier rounds of NELS:88 and HS&B. Inability to pretest all aspects of the main study design and procedures therefore seemed an acceptable price to pay for the opportunity to investigate the potential contribution of free response items.

Upon completion of field test data collection, NORC and ETS conducted an evaluation of the instruments and procedures. Typically, these data are used to inform planning for the main study regarding procedures to be used or refined. Additionally, such analyses serve to identify means to improve the measurement properties of test and questionnaire items, as well as to identify items which need to be deleted or modified for instrument length or item format reasons. This abstract summarizes the field test study design and data collection procedures used and highlights the recommendations for changes in the instruments and procedures.

### Design of the Field Test

The Second Follow-Up Field Test included four major survey components: students, dropouts, parents, and school administrators. (The Teacher Survey, funded for the main study, was not included in the field test.) As noted above, the student and dropout sample members were those initially selected for the Base Year Field Test and those added through sample freshening procedures in the First Follow-Up Field Test. The majority of the sample members still in school were in the last half of their senior year in high school. Some sample members had withdrawn from school. Students completed a questionnaire and cognitive tests in two subject areas during in-school group sessions. Although in the main study students have traditionally completed all four tests in the cognitive battery, as in previous NELS:88 field tests, 2FU field test participants were tested in two subject areas only, so that more items could be assessed to produce reliable and valid test items in the main study. The cognitive tests included multiple choice and free response items. Dropouts were identified and invited to group administrations; those participating completed a questionnaire and cognitive test.

Since the methodology for the parent survey is fundamentally the same as that used in HS&B and NELS:88 Base Year, the major purpose of the parent component field test was to gather sufficient observations for evaluating the performance of questionnaire items. A subsample of field test schools was therefore selected to pretest the parent questionnaire and mailing materials. The Base Year respondent or the parent most knowledgeable about the teen's education was targeted.

The school principal of each participating field test school was asked to complete a School Administrator Questionnaire; principals could designate other school staff to complete the first four sections of the instrument, reducing burden on principals' time. A field test of course offerings and enrollment data collection procedures is currently ongoing; a similar field test for the transcript component will begin in fall 1991.

## Data Collection Procedures

Pre-data collection activities included instrument development (see next section), field staff recruiting and training, tracing sample members, and securing the cooperation of the five field test states and the districts, and schools attended by sample members. During the fall of 1990, sample members were traced by telephone contact with the school attended in the tenth grade; other possible sources of information (e.g., household, school where most tenth grade classmates attended) were pursued, with minimal dependence on "field tracing" (following leads in person). These tracing procedures were effective -- ninety-five percent of sample members were successfully located, most (86 percent) at their tenth grade school. (While for purposes of the field test, a 95 percent level of tracing was sufficient, in the main study we plan to trace over 99 percent of the sample, as was done in the prior round.)

To secure chief state school officers' permission to contact districts within the state, letters explaining the study were sent; NORC project staff followed up with a telephone call to answer questions and discuss the state's participation. When the state's cooperation was obtained, a letter and NELS:88 informational materials were sent to district superintendents. NORC field managers contacted district superintendents to seek permission to contact principal(s) of schools in the district attended by NELS:88 students. Ninety-two percent of districts contacted granted permission to contact school principals.

School principals were asked to host a NELS:88 Second Follow-Up Survey Day, or data collection session, to be attended by sample members and sufficient other seniors to total 35-40 students. Visits were made in the fall to each school by NORC interviewers to freshen and augment the sample, confirm student home addresses, and make initial arrangements for data collection.

Data were collected from students between February 4 and March 23, 1991. A total of 2,254 students in 94 schools participated in the 2FU Field Test, for an overall response rate of 70.5 percent. Longitudinal cohort members and augmented students exposed to NELS:88 for the first time participated at similar rates (69.9 percent and 70.9 percent respectively), but students added through freshening procedures in the 1FU and 2FU were less inclined to participate (62.5 percent). Field test resources were used to quickly and efficiently amass the number of observations required for the cognitive test assessments, rather than to maximize response rates. Even though more time, and resources, will be available for increasing response rates in the main study, special note should be made of traditional difficulties with securing the participation of twelfth graders. It cannot be forgotten that the quality of performance on the data collection task will affect the overall success of the Second Follow-Up more than any other single factor. Main study procedures will be adjusted to take into account the disengagement of second term seniors and any resulting disinclination to participate, with special attention to sample members added through freshening.

Many school coordinators complained about the level of burden imposed upon them; we recommend reinstating the honorarium to partially compensate them and to serve as a token of appreciation for their contribution to the project. We also suggest several ways that the school coordinator's responsibilities can be reduced. The fullest support of the school coordinator is also required if student response rates are to be maximized.

A total of 68 dropouts (all cohort members) were identified during the autumn tracing activities; another 40 (combined cohort and augmented) were identified during the data collection period. Field test resources were focused on maximizing the effectiveness of group administration participation procedures, since these have not been applied previously in NELS:88. Owing to the small size of the field test dropout sample, special emphasis was put on qualitative means of assessing the dropout questionnaire. Ten cognitive interviews were completed with dropouts in the Chicago and Washington, D.C. areas for qualitative data to evaluate the instrument. A special issue that must be confronted in the Second Follow-Up is that dropouts were somewhat differently defined in NELS:88 1FU and HS&B 1FU. NELS:88 2FU should strive to maintain distinctions necessary to ensure the comparability of its dropout statistics to both studies, and should carefully review basic definitional choices and their implications.

Of the 688 parents in the field test subsample, data were obtained from 506 for a 73.5 percent response rate. Questionnaires were mailed to all subsampled parents on February 12th. Prompting calls were made and telephone interviews were conducted during the period from March 4th and April 8th.

School Administrator Questionnaires were distributed through the school coordinator, (except to the 31 schools with Survey Days scheduled in the first weeks of data collection. The instrument was delayed at the printer and questionnaires were mailed directly to the principal in these schools.) Interviewers picked up questionnaires completed by their visit to the school for Survey Day and prompted nonresponding principals near the end of the data collection period. Sixty-five schools submitted completed questionnaires.

### The Instruments

Survey questionnaires were developed during the summer of 1990 by NORC staff in collaboration with NCES personnel and representatives of the educational policy and research communities. The major study goals and design constraints which guided the development of the field test instruments included: 1.) limitations on the number of questionnaire and test items that can be asked; 2.) the general purpose nature of NELS:88 3.) the defining longitudinal purposes of NELS:88; 4.) the desirability of using NELS:88 data in cross-cohort analyses with data from prior education longitudinal studies (the National Longitudinal Study of the High School Class of 1972 [NLS-72] and High School and Beyond [HS&B]); and 5.) the strong need for integration and articulation between component questionnaires and across rounds of NELS:88 data collection.

This report contains a comprehensive methodological evaluation of the questionnaires. These analyses of questionnaire data include reliability indices for scalable elements, checks of inter-item consistency, examination of frequency distributions, and analyses of item nonresponse. In addition to analyses of individual items and scales, a number of questionnaire experiments--concerning item wording, item response format, and item context--were included to identify and provide a basis for controlling sources of nonsampling error. "Cognitive laboratory" techniques have also been utilized in pretesting and evaluating field test items. Results contained herein, and other inferences made from the field test data, should be viewed within the limitations of the samples. Recommendations are given primarily from a methodological perspective, in anticipation of further review and input from data users in the next phase of the instrument development process. Information provided in Appendix E may facilitate such further interpretation of the analytic utility and measurement potential of field test questionnaire items.

Four questionnaires and five cognitive tests were developed for the field test. The **Student Questionnaire** included many content areas that overlap with the previous waves of NELS:88. Questions were added to obtain information about the transition of the noncollege bound to the labor force, and of others to colleges and other postsecondary institutions. By design (it is desirable to test a large item pool in the field test) the student questionnaire exceeded the target length for the main study. The Second Follow-Up Field Test experience suggests that the instrument should be shortened to about 65-70 percent of its present length, or to a total of approximately 90 questions. We also recommend simplification of item formats, and of instructions and question wording. Results of experiments included in the student questionnaire indicate that response categories in behavioral frequency questions should be revised to minimize bias, that item order had no effect on either evaluations of the school or reports of school-related behavior, and that data quality can be improved through the use of explicit affirmative or negative response categories rather than the "mark all that apply" format.

The **Not Currently In School** (or Dropout) **Questionnaire** was modified to obtain data for the analysis of the different experiences and processes involved in early and late dropping out. Another objective in the development of the instrument was maximizing the articulation between the Dropout and the Student Questionnaires. Through the qualitative assessment of cognitive interview data, First Follow-Up item nonresponse data, and the few questionnaires completed during the 2FU Field Test data collection period, we offer recommendations for simplified question wording and formatting, as well as identifying several sections of the questionnaire that would benefit from further refinements and iterative qualitative assessments.

The **Parent Questionnaire** repeats key intra- and cross-cohort comparison variables from the NELS:88 Base Year and HS&B. While the instrument contains questions to examine the home educational support system, new items have been added to elicit information about planning for the sample member's entry into the labor force or postsecondary education. Our recommendations from field test experience include shortening the instrument and revising the format of some types of questions. Results of experiments included in the Parent Questionnaire indicate that item order can affect parents' reports of satisfaction with their child's education--leading to a recommendation to place the general satisfaction item before any eliciting more specific judgments--and that parental reports of drug or alcohol use problems are also affected by item order. To minimize the social desirability bias, we recommend either employing reports of drug/alcohol problems among the teen's friends as a proxy measure for the teen's parent-reported behavior, or placing the question about friends' use before the question about the teen. Relevant to the design and planning of the 2FU main study, we additionally recommend consideration of a supplement to the Parent Questionnaire to be completed by Base Year nonrespondents, and parents of sample members who have been added to the cohort through sample freshening procedures in the 1FU and 2FU.

The **School Administrator Questionnaire** was designed to gather descriptive data on school policies and practices. One priority in development of the 2FU instrument was to reduce significantly the number of questions asked of school personnel; hence, the size of the questionnaire has been reduced to approximately one-half the number of data elements used in NELS:88 1FU. (Excessive length was a problem for almost all the 1FU questionnaires; a goal of the NELS:88 2FU should be to try to more nearly approach the questionnaire length targets of preferably less than but certainly no more than 30-45 minutes, as prescribed in NCES Standard 87-03-01). While complaints about excessive burden were received from a few school principals during the 2FU Field Test, we anticipate no unsurmountable problems in the main study if the length of the School Administrator Questionnaire does not increase. We do recommend consideration of the addition of a series of supplemental questions for schools new to NELS:88.

**The Cognitive Test Battery.** The primary objective of the NELS:88 cognitive test battery is to measure cognitive growth over time--specifically, between eighth and twelfth grade. An additional objective of the NELS:88 2FU test battery is to provide a maximally reliable cross-sectional attainment measurement in grade 12, thus supplying a benchmark for postsecondary rounds of NELS:88. A major question to be answered by the field test was the extent to which free response items can contribute to these goals. (Such items were not included in the Base Year or 1FU, which used multiple choice formats exclusively.) The field test report analyzes test item response rates and patterns, the contribution of multiple choice and free response items to overall reliabilities, and the general psychometric properties of the free response items included in the field test cognitive battery. These analyses suggest that the Base Year and First Follow-Up test design should be retained in the 2FU, if the measurement of longitudinal gain is to take place within the design and logistical constraints of NELS:88. We do recommend, however, that NCES give serious consideration to the possibility that the free response item database constructed in the field test be made available to other potential users in the psychometric community.

### Special Issues Pertaining to the Main Study

Not all elements of the Second Follow-Up Main Study were included in the Field Test. Given the multiple, inter-related research goals of NELS:88, it is imperative that planning for the main study take account of all components. This report therefore also summarizes several key issues related to the design of main study components that had no counterpart in the field test. These components are the school effects supplement, the study of base year ineligibles, the teacher survey, and the early graduate supplement.

NELS:88 is a study of both individual transitions and of school effects upon individual outcomes. NELS:88's ability to study institutional effectiveness is enhanced by inclusion in the study of a substantial school effects supplement designed to provide robust and representative within-school samples of students within a subsample of NELS:88 high schools. Plans for the school effects supplement in the NELS:88 2FU are outlined in Chapter 11. Also included in the chapter is a discussion of issues relating to a special followback study of sample members determined to be ineligible in the Base Year. The study of these excluded students is necessary if NELS:88 is to produce true national estimates, such as a cohort dropout rate. In addition, since eligibility status can change over time, a longitudinal survey that will generate representative cross-sections for future time points must both freshen its student sample (to give a chance of selection to 1990 tenth graders and 1992 twelfth graders who were not eighth graders in 1988), and to give baseline ineligibles whose status has changed some chance of selection into later rounds of the study. Finally, chapter 11 describes current plans for the Teacher Survey and for the Early Graduate Supplement questions.

## INTRODUCTION

This introduction is a brief summary and contains: an historical perspective of the National Education Longitudinal Study of 1988 (NELS:88), the components and instruments included in the Second Follow-Up Field Test and Main Study designs, the schedule of project activities, the goals and guiding principles of the Field Test, and the organization of this report.

The major portion of the report is dedicated to a methodological examination of the data from the field test study. Standard techniques, such as examining nonresponse and filter/dependent consistency, were used along with innovative techniques, such as question experiments and cognitive interviews, to evaluate questions in the student, parent, dropout, and school questionnaires. The analysis provides a comprehensive methodological evaluation of the questionnaires within the limitations of the properties of the field test samples. Individuals wishing to make assessments concerning utility of the questions for substantive analysis may be aided by the question frequencies presented in Appendix E.

**Background:** With the award of the Base Year contract in February 1986, NELS:88 joined the National Longitudinal Study of the High School Class of 1972 (NLS-72) and High School and Beyond (HS&B) as the third in a series of longitudinal studies sponsored by the National Center for Education Statistics (NCES) of the U.S. Department of Education. These studies are designed to provide trend data about critical transitions experienced by young people as they develop, attend school, and embark on their careers. Given the challenges facing America's schools -- to educate all our young people for the next decade, regardless of family circumstances -- NELS:88 will complement and strengthen state and local efforts by furnishing new information on how school policies, teacher practices, and family involvement affect student educational outcomes (academic achievement, persistence in school, and participation in postsecondary education). NLS-72 and HS&B surveyed high school seniors (and sophomores for HS&B) through high school, postsecondary education, and work and family formation experiences. Taken together, these three longitudinal studies provide not only measures of educational attainment but also rich explanations of the reasons for and consequences of academic success and failure. Both NLS-72 and HS&B have influenced the school reform movement; NELS:88 will provide comprehensive data for gauging the degree of the reform movement's success.

NELS:88 is the first national longitudinal education study to begin surveying students as early as the eighth grade. In the Spring of 1988, base year data were collected from a sample of over 24,500 eighth grade students attending over 1,000 schools (800 public and 200 private schools including parochial institutions) across the nation. The students' parents, their teachers, and their school principals were also surveyed. The First Follow-Up revisited over 18,000 of these same students in 1990, when they were in the tenth grade. Four questionnaires -- student, school administrator, teacher, and dropout -- were administered. As in the Base Year, cognitive tests in four achievement areas (reading, mathematics, science, and social studies) were given to cohort members.

**The Design and Schedule of the Second Follow-Up Field Test:** -- The Field Test phase of the Second Follow-Up of the National Education Longitudinal Study of 1988 (NELS:88 Second Follow-Up or 2FU) began upon contract award on May 25, 1990 and culminates with the submission of this report. The NELS:88 2FU Field Test -- with the objective of evaluating the effectiveness of questionnaires, cognitive tests, and survey procedures to be used in the 1992 2FU Main Study -- was conducted at selected schools in California, Florida, Illinois, New York, and Texas with the same approximately 900 students who participated in the 1FU Field Test and approximately 2,000 12th graders added to the Field Test sample to allow for assessment of free response cognitive test items. Questionnaire development activities were concentrated in the summer months of 1990. Preparations for data collection started in the late summer and continued through December. Data were collected from students, dropouts, and school administrators between February 4 and March 23, 1991. The data collection period for the parent component was initiated by a mailout of questionnaires on February 12th and continued through April 8. A total of 2,254 students in 94 schools participated, completing a self-administered questionnaire and a cognitive test battery which included free response items; a subsample of 688 parents was selected and of these, 506 provided data. A total of 29 school leavers completed the Dropout Questionnaire and 65 school administrators filled out the school questionnaire. This report describes and evaluates the procedures and instruments used to survey these respondents and recommends changes for implementation in the Second Follow-Up Main Study.

**The Design and Schedule of the Second Follow-Up Main Study:** The design for the NELS:88 2FU main study to be conducted during the 1992 school year includes five questionnaires -- student (including supplemental questions for early graduates), dropout, parent, teacher (not assessed in the Field Test), and school administrator. The first phase of preparatory activities -- including tracing cohort members and contacting districts and schools for cooperation in the 2FU Main Study -- began in February and ended in mid June; a second phase is scheduled for this fall and will consist of sample freshening, the identification of eligible teachers, making initial arrangements for data collection, and additional tracing.

The data collection period, or third phase, is scheduled for the spring of 1992 between February and June. During that time NORC representatives will administer the Student Questionnaire and Cognitive Test to NELS:88 sample members in a group setting at schools with sufficiently large cluster sizes. (School coordinators will be asked to administer sessions in smaller cluster schools.) The students' parents, teacher(s), and school principal will be asked to complete a self-administered questionnaire. Schools will also be asked to provide course offerings and enrollment data and copies of academic transcripts for students who have authorized transcript release. Data collection from parents is scheduled for mid May through July. The course offering request will be made during preparations for Survey Day; course enrollment data will be obtained during spring and summer of 1992. Transcript acquisition will not begin until fall of 1992.

**The Student Questionnaire**, now geared toward high school experiences, assesses both students' school-related and personal experiences. Specific content areas will collect data regarding academic achievement, student perceptions and feelings about their curriculum and school, family structure and environment, social relations, aspirations, goals, attitudes, and values, particularly as they relate to school and occupational plans. The National Science Foundation has sponsored a mathematics and science item supplement. The Office of Bilingual Education and Minority Language Affairs (OBEMLA) has sponsored a language minority supplement. The Student Questionnaire also contains items never asked before in previous longitudinal education surveys. For example, the NELS:88 Student Questionnaire contains items on family decision-making structure. Such items will aid in an understanding of how decisions concerning students' education which are made during the critical transition from secondary school to post-secondary education or the work environment affect students' academic achievements, and future education and/or occupational plans. The Student Questionnaire is allotted 50 to 60 minutes to administer. (See Section 3.9 of this report for detailed recommendations.)

During the student survey session, students will be asked to complete a self-administered questionnaire and then take a timed cognitive test battery. The **Cognitive Test**, designed by the Educational Testing Service (ETS), consists of curriculum-sensitive items for measuring cognitive growth between eighth and twelfth grades in four achievement areas: reading, mathematics, science, and social studies. The cognitive test administration is allotted 85 minutes of the session. At the time of this report submission, no decision has been made about inclusion of free response items in the main study cognitive test battery. (However, on the basis of field test results, NORC and ETS have recommended that free response items not be included in the 1992 battery. See Section 3.10.5 for detailed recommendations.)

**The Dropout Questionnaire** will permit investigation of the factors responsible for students dropping out of school. The questionnaire will be administered in a group setting at an off-campus central site, or in a small number of cases, in the respondent's home or by telephone or mail. (See Sections 4.5 and 5.4 for detailed recommendations.)

A series of supplemental questions for **Early Graduates** will be included in the Student Questionnaire. At the time of the 2FU, it is expected that some members of the 1988 eighth-grade cohort will have graduated from high school. Based on our experience with early graduates in HS&B, we expect that there will be about 950 early graduates in the 2FU. The intent of the supplement is to document the reasons for and circumstances of early graduation, the adjustments required to finish early, and respondents' activities compared with those of other out of school survey members such as dropouts. Each early graduate will complete this special Early Graduate Supplement, the Student Quesitonnaire, and a battery of cognitive tests in an off-campus (usually group) administration. Content areas to be covered in the brief supplement will include: reasons for graduating early, when decision was made (what grade), persons involved in the decision, course adjustments required, school requirements, and post-secondary education and work experience. (See Section 11.4 for detailed recommendations.)

The school principal is the designated respondent for the **School Administrator Questionnaire.** The questionnaire is designed to gather descriptive information regarding the school's teaching staff, student body, and school policies and offerings. Additional questions will inquire about transition patterns, enrollment, student demographic and linguistic characteristics, school tracking policies, extra-curricular programs, school facilities, participation in federal assistance programs, staff to student ratios, patterns of dropping out, school climate, faculty size and characteristics, school rules and disciplinary programs, and special features or problems of the school. OBEMLA has sponsored a series of language minority items. The data gathered from the School Administrator Questionnaire will be available for analysis of the learning environment and experiences of the sampled students. The questionnaire is designed in such a way that an administrative assistant or other knowledgeable school official can complete sections 1 through 4, but the final section, section 5, must be completed by the principal. This approach to the instrument's design, and a much shorter instrument are anticipated to reduce the burden for the principals considerably. (See Section 9.5 of this report for detailed recommendations.)

The major purpose of the **Teacher Survey**, sponsored by the National Science Foundation (NSF), is to obtain information about NELS:88 sample students' school-related characteristics and experiences, especially those which are believed to affect students' educational development in mathematics and science. The Teacher Questionnaire will collect data about school and teacher characteristics (including teacher qualifications and experience), and classroom teaching practices. There are two components to the Teacher Survey currently under discussion: 1) The National Science Foundation has sponsored a component to collect data from **teachers of mathematics and science** who have NELS:88 sample member(s) in their classes; and 2) NCES has funded a component to collect data from teachers of mathematics, science, English, and social studies in **School Effects Supplement (SES)** schools who have NELS:88 sample member(s) in their classes. (See Section 11.3 of this report for detailed recommendations.)

The object of the supplementary **Parent Questionnaire** is to collect reliable and valid information on each student's family background, the parent's relationship to the student, and the parent's knowledge of, and involvement in, the student's family, school, and community life, and future plans. OBEMLA has sponsored a language minority supplement for the questionnaire. Parent responses provide data for the analysis of student behaviors and outcomes. The 2FU Main Study will target one parent (or guardian) for each eligible NELS:88 student. Through the use of cover letters and other printed materials, efforts will be made to direct the Parent Questionnaire to the parent or guardian who completed the Base Year questionnaire or who knows most about the student's educational activities and related behaviors. (See Section 7.8 for detailed recommendations.)

For each high school selected to participate in the Second Follow-Up survey, **course offerings, enrollment, and transcript data** will be collected, and merged with data collected by the school questionnaire. This information will be available at the student level to allow for analysis of student course-taking patterns and these data will enable researchers to classify schools as to the emphasis they place on selected curricular areas (for example, mathematics, science, foreign language), and will provide a mechanism for identification of schools that emphasize college placement (for example, high percentage of students enrolled in advanced mathematics) or that emphasize vocational preparation. Course offerings data will also be used to code student transcripts. (See Chapter 10 for further details and recommendations.)

**Objectives and Principles of the Field Test:** The single greatest driving force behind the design of the Second Follow-Up Field Test was the inclusion of free response (or open ended) items in the cognitive test battery. The decision to test these for possible application in the main study had numerous implications for the design of many other field test components. Several examples are discussed below.

The original design of the 2FU Field Test required data collection from most members of the field test cohort (n = 1,086) for assessment of instruments and procedures. However, the decision to test free response items meant that a significant number of additional observations would have to be obtained. Rather than selecting more schools to participate in the field test, the sample of students within each field test school was augmented through a random selection to ensure a minimum of 2,000 completions. This approach, while resulting in the same outcome (the necessary number of tests for analysis), was judged as more likely to be time and cost efficient.

Knowing that young people are more difficult to gain participation from, and that second-semester seniors are perhaps the most difficult, we sought permission from each field test school to augment the total sample to 40 seniors. The invitation to participate to these much larger groups ensured the completion of 2,000 tests, without driving up costs through Make Up Day sessions and other efforts to obtain data from nonresponders.

The large cluster sizes in most schools, however, limited the activities that could be requested of school staff. For example, the need to invite so many students from each school resulted in excluding the pretesting of school coordinator administered data collections session (except in a handful of cases). Such sessions were conducted in the First Follow-Up, and in ideal circumstances Second Follow-Up data collection procedures would have been tested. An assessment could have been made about the likelihood of school coordinators conducting the sessions according to 2FU procedures and the effect of no symbolic compensation (The school coordinator honorarium used in HS&B and prior rounds of NELS:88 was not included in the approved plan for the 2FU.) on school coordinator's motivation and behavior in the more burdensome scenario. Refinements could then have been made before the main study based on more than anecdotal data.

With the number of cognitive test observations driving the design, there are significant aspects of the study that could not be assessed (for example, school cooperation and propensity of students to participate), even if the longitudinal cohort and augmentation are separated for analysis. The inclusion of the additional students irreparably alters reactions to the study in ways that cannot be controlled for statistically. Fortunately, many of the procedures to be employed in the 2FU have been well tested in prior waves of NELS:88 and HS&B; the 2FU Field Test sought only to define ways to adjust these to the changing environment -- more spartan school budgets, understaffing at many schools, and overcrowding in schools in some areas of the country -- and to refine the procedures for greater cost efficiency. A few procedures (for example, tracing sample members by telephone and collecting data from a majority of dropouts in group administrations) had not been applied to NELS:88 previously and were targeted for a more thorough assessment. The Second Follow-Up Field Test was designed and implemented with full knowledge that some procedures were not being tested and that assessments of certain procedures and outcomes would have to be made carefully, taking the necessary differences between the field test and main study into account to the greatest extent possible.

**Organization of this Report:** This report contains eleven chapters and five appendices. The first chapter describes preparations for the collection of field test data. The second chapter describes the content and objectives of the Student Questionnaire, the experiments implemented in the instrument, and the procedures employed in collecting, preparing, and processing data from students in the in-school sessions; the third chapter presents the results of our analysis of those data. The fourth chapter describes the objectives of the Dropout Questionnaire, and the data collection and preparation procedures used to obtain and process the dropout data. Chapter five discusses the qualitative data collection procedures used to assess the Dropout Questionnaire and the findings of those cognitive interviews. Chapter six describes the content and objectives of the Parent Questionnaire, the experiments implemented, and data collection and preparation procedures; chapter seven presents the results of analysis of those data. Chapter eight describes the content and objectives of the School Administrator Questionnaire and data collection and preparation procedures; chapter nine discusses findings based on our analyses. Chapter ten describes the on-going pretesting of course offerings and enrollment data collection procedures. Chapter eleven summarizes our recommendations and discusses our plans for the Early Graduate supplemental questions, the Teacher Survey, and several issues related to the follow-up of ineligible sample members and the School Effects Supplement.

Four appendices are included with this report in Volume I. Appendix A contains minutes and correspondence documenting the meetings held in 1990 regarding instrument development. Appendix B contains crosstabulations that support questionnaire experiment results reported in Chapter 3. Appendix C includes tables that support the analysis of the cognitive test data. Appendix D contains summaries of the cognitive interviews conducted to qualitatively assess the Dropout Questionnaire. Volume II contains Appendix E, which includes each of the four questionnaires and the data collected (frequencies, and for the Student Questionnaire, retrieval rates for critical items).

Readers may find it helpful to examine the question frequencies listed in Appendix E. A guide to reading the frequencies is included at the beginning of that appendix. When examining the frequencies, it is important to keep in mind that though the field test sample exhibits heterogeneity, it is not a probability sample from which population estimates can be made. Also, for two of the questionnaires (School and Dropout), the sample sizes are very small. The small sample sizes should be taken into account when considering differences in percentages across response categories.

## Chapter 1: Preparation for the Field Test

Upon award of the Second Follow-Up contract on May 25, 1990, NORC and ETS personnel began their background review of materials relating to the overall study design, methodologies and instruments. Several lines of inquiry were pursued, among them: the examination of prior studies (particularly HS&B and NLS-72), the review of policy issues to be addressed by the instruments, and inquiry into parallel lines of research. Taking the item specifications in the Request for Proposal as a starting point, the intent of the background review was to identify items, themes and methods employed in the past that would be valuable for the NELS:88 Second Follow-Up and (in the case of NLS-72 and HS&B) that could provide cross-cohort comparability in key measures and methods; and to identify appropriate new content areas and items for the questionnaires and cognitive tests. The quest for appropriate new items entailed canvassing the item banks of major national studies. But it also required that we examine more local but intensive case studies, and in some instances consider the research hypotheses of the policy research literature to discover and address basic gaps in national data sources.

This phase of the study involved discussions with staff of research centers, conversations with officers of specialized educational constituencies, and meetings with various policy offices within the U.S. Department of Education and other public agencies. This phase of the study also involved an extensive literature review.

In tandem with the background review, a series of planning meetings was held, involving both ad hoc bodies and the NELS:88 Technical Review Panel (formerly the National Advisory Panel). The first of these sessions, the preliminary Content Outline Meeting, took place at NCES on June 4, 1990, and solicited the reaction of participants to draft content outlines for the student, parent, school and teacher questionnaires. Input was sought in the following areas:

- New content domains that would help inform issues related to the nation's education goals,
- Content domains in which pertinent research is now under way,
- Content domains overlooked or under-represented in the current content outlines, and
- Assessment of the need for additional contextual information on schools, communities and families.

(See Appendix A for a detailed summary of the proceedings.)

The NELS:88 Technical Review Panel (TRP) met June 26-27, 1990, at the NCES offices in Washington, D.C. The purpose of the June TRP meeting was to come to closure on the content areas to be covered in the Second Follow-Up field test questionnaires. In addition, panelists were informed of Base Year products; the progress of the First Follow-Up; and issues pertaining to the cognitive test battery. The agenda for this meeting is reproduced in Appendix A.

A planning meeting on NELS:88 cognitive tests was held in Boulder, Colorado, on June 13, 1990. The purpose of the meeting was to discuss the following topics:

1. Test and item specifications
2. Proposed subscales and proficiency levels
3. Common linking items in the G-8, -10, and -12 tests
4. Plans for comparisons with HS&B

5. Linking topic coverage/"opportunity to learn" items with cognitive test items and coursework
6. Options for open-ended items in the 1991 field test
7. Possibility of impact of contextual effects and item format on student responses
8. Possibility of equating NELS:88 and NAEP

(A summary of the Boulder meeting appears in Appendix A.)

Finally, on August 13-14, 1990, a second meeting was held on NELS:88 cognitive testing issues, in Washington, D.C. at the offices of NCES. The purpose of the meeting was to further explore options for the field test, particularly in connection with the field-testing of free response (open-ended) test items. Topics discussed included: specifications and objectives for open-ended items; models of the learning process, considered as an assessment framework; scoring methods; various format, content, and method of administration options for open-ended items; and prospects of equating NELS:88 and NAEP mathematics scores. Correspondence summarizing conclusions of the August test meeting can be found in Appendix A.

As a result of the initial review, and the input received at the various meetings, basic questionnaire drafts were constructed, developing and elaborating the questionnaire specifications originally provided by NCES. Items were developed iteratively, with successive drafts reviewed by NCES, OBEMLA, Technical Review Panel members, and consultants.

Thus far we have outlined the process by which NCES and contractor staff sought to define appropriate content coverage. However, a second critical dimension of the field test instrument development process was to ensure that items on the NELS:88 instruments elicit data of the highest statistical quality. While the field test data analysis plan (to be discussed later in this report) was devised to realize this objective, the need for assurance of accuracy of measurement played a role in the developmental process as well. It did so in two ways. First, item selection was guided by what was known about the psychometric properties of prospective questionnaire items. If a question had been used in another study, for example, frequencies distributions could be inspected to gauge whether there had been useful variation, and item nonresponse, reliability (for scalable questionnaire components), and interitem consistency information might also be available. Second, the process of refining items took heed of the methodological literature on item writing and questionnaire design, and focus groups and pretests were used to ensure that clarity of meaning was achieved and that response categories were appropriate.

The NCES Interdivisional Review Panel reviewed the draft instruments August 8th and 9th, 1990. Item justifications were written in the course of August and a full Office of Management and Budget (OMB) forms clearance package submitted to the Department of Education's Information Management and Compliance Division (IMCD) September 17, 1990. The final version of the forms clearance package was delivered to OMB on October 16, 1990. Revisions to the questionnaire were made in accordance with OMB recommendations and final clearance was received December 27, 1990. The questionnaires were then formatted and prepared for printing, in anticipation of the start of the field period in the first week of February, 1991.

In addition to development of the instruments, there were many other activities in preparation for field test data collection. Field staff were recruited, hired, and trained; the sample was freshened and augmented; field test cohort members were located and their status as a student or a dropout was confirmed; and participation in the study was sought from states, districts, and schools. These topics are covered in detail in Sections 1.2 through 1.7 of this chapter.

## 1.1 Development of Field Test Instruments

The general approach to field test questionnaire development and test construction was shaped by the goals and constraints imposed by the overall design of NELS:88. The following study goals and design constraints are particularly important. Each is described in detail below.

1. **The number of questionnaire and test items that can be asked is severely limited;**

2. **NELS:88 is a general purpose data set designed to serve a broad constituency and to fulfill a variety of purposes;**

3. **The primary research objectives of NELS:88 are longitudinal in nature;**

4. **NELS:88 must also support cross-cohort comparisons with NLS-72 and HS&B;**

5. **Articulation -- between questionnaires and across rounds of data collection -- is a paramount requirement of NELS:88 instrument design.**

1.) **The number of questionnaire and test items that can be asked is severely limited.** The maximum time that can be allotted for the cognitive tests is 85 minutes. Time available for the student questionnaire is one hour; a 60 minute time limit presupposes that the average student will complete the instrument in about 45 minutes. Clearly, there are far more items, and many more worthy themes, than can be accommodated in this brief compass. However, given the need to test alternative versions of items in the field test, and given the desirability of maximizing the size of the field test item pool so that more successful data elements could be retained and less successful data elements dropped, it was agreed that the target length for field test questionnaires could exceed the intended main study questionnaire length by up to 15-20 percent.

2.) **NELS:88 is a general purpose dataset designed to serve a broad constituency and to fulfill a variety of needs.** First and foremost, the purpose of NELS:88 is to inform educational policy at all levels of government and administration. Given resource limitations, highest priority must be placed upon collecting policy-relevant data on educational processes and outcomes. Because NELS:88 must serve multiple, potentially competing broad purposes, a dominant feature of the instrument development process is the need to constantly prioritize questionnaire goals as well as find new ways to harmonize conflicting demands, all the while striving to maintain a coherent focus for these broadly-based instruments.

A general purpose dataset such as NELS:88 must meet the needs both of the policy community (which must evaluate programs and make policy decisions) and the research community (which needs data to build and test causal models of the educational process, and which itself represents multiple social science disciplines and research paradigms). Such a dataset must serve descriptive statistical reporting purposes as well, providing a basis for characterizing the education system in its various aspects. (Thus, for example, NELS:88 First and Second Follow-Ups must both provide national estimates of school-leaving -- a cohort dropout rate -- as well as data for multivariate analyses of the dynamics of dropout behavior). The perspectives of statistical agencies, and of policy and basic researchers, impose somewhat different statistical design and content demands on the study; a major challenge of the NELS:88 Second Follow-Up is to maximally accommodate divergent needs within the severe constraints of resources, testing and surveying time, and reasonable respondent burden.

A further manifestation of NELS:88's role as a general purpose dataset is reflected in the tension between comprehensiveness of content and specificity or depth of measurement that must be confronted in the questionnaire design process. This challenge may be faced and met in several ways.

- One may look for points of intersection, where distinct goals that might otherwise conflict may be operationalized in a manner that proves complementary.
- One can strive to maximize the amount of information collected by the survey, by carefully planning the sequencing and articulation of content across survey waves and across various survey populations within a given wave.
- And one must be prepared to prioritize, in recognition that a general purpose survey that must sustain great breadth of content will be able to measure but few things in depth.

In general, prioritization will favor the longitudinal objectives of the study, and its capacity to uniquely illuminate educational processes that affect cognitive growth in high school.

3.) **The primary research objectives of NELS:88 are longitudinal in nature.** The study provides the basis for within-cohort comparison by following the same individuals over time in order to measure cognitive growth and record key transitions; a longitudinal design permits researchers to make causal inferences to a degree not possible in cross-sectional surveys. Priority, therefore, from the very beginning of the study, has been placed on items expected to be most useful for predicting or explaining future individual- and group-level cognitive, behavioral and affective outcomes, and on repeating key change measures from prior rounds. The study's longitudinal objectives drive complementary needs for content stability and for content change.

A core set of items must be maintained so that researchers can examine changes in individuals over time. From a cross-wave perspective, the requirement is to measure the same thing the same way repeatedly -- repeated measures of the same individual must, to be valid and reliable, involve a high degree of content stability in the measures employed. At the same time, each new critical transition encountered by a longitudinal survey requires that new benchmark data be obtained.

In a longitudinal study such as NELS:88, the respondent population, as it moves through adolescence, is rapidly changing, and questionnaire content must mirror these changes in the external situation and internal development of respondents. But so too is the state of educational theory developing. New items are necessary both to accommodate changes in age, stage, and situation of the sample, and as well the advancement of the field. The instruments must be aligned with contemporary theory, or they will have no chance of capturing so-called "emerging issues."

The Second Follow-Up will mark the second major transition to be charted by NELS:88. The eighth grade survey in 1988 obtained measures at a point just prior to entry to secondary school. The 1992 survey will obtain a third set of achievement scores by means of the cognitive test battery, secure secondary school transcripts, and obtain further student questionnaire and contextual data, at a point that for most respondents will represent the completion of high school and a point where its cumulative impact may be measured. But of course the end of high school ushers in a new beginning as the critical transition from high school completion to labor force entry, postsecondary education, and family formation unfolds. Thus the Second Follow-Up becomes a further baseline for posing new questions that have become germane with the passage of time and advancement of the cohort into a new developmental or life stage.

These new content areas for the Second Follow-Up embrace, for example, the influence of teachers, family, peers and counselors on basic occupational and postsecondary educational choices; the degree of family provision for postsecondary education; and the influence of the availability of resources on the educational aspirations and expectations of sample members. In addition to gathering outcome data that can be related to precursor events measured in prior rounds, the Second Follow-Up questionnaires must lay the foundation for studying such basic postsecondary education issues as undergraduate access and choice, persistence, progress through the curriculum, attainment and outcomes, and rate of individual and societal return on postsecondary education; transition to the world of work; prospects of return to school or alternative completion for those who have dropped out of secondary school; and the manner in which sample members form and realize their goals in education, employment and family formation.

Because the Second Follow-Up is the culmination of the process of measuring the impact of secondary schooling, items from prior waves will predominate. Nevertheless, substantial numbers of new items, benchmarking the transition to the postsecondary world, will be asked, and will be critically important to the Second Follow-Up dataset. Continuing this move to new items and domains, in the Third Follow-Up, when the cohort is largely outside the secondary school system, the number of repeated items is expected to decrease drastically. High School and Beyond offers a probable precedent. About two-thirds of the sophomore cohort student questionnaire items were essentially the same between tenth and twelfth grade; but three-quarters of the items on the HS&B Second Follow-Up (1984) questionnaire were basically dissimilar to the baseline (1980) items.

To say that NELS:88 is driven by longitudinal research objectives, is not, however, to say that cross-sectional objectives are unimportant. The NELS:88 sample is freshened (for an explanation of sample freshening, see Section 1.3.2.) in each of the in-school follow-ups to provide three nationally representative cross-sections -- eighth graders in 1988, tenth graders in 1990, and twelfth graders in 1992. Therefore, a further guideline for field test instrument development was that where consistent with the longitudinal aims of the study, the Second Follow-Up instruments should also seek valuable cross-sectional data.

4.) **The Second Follow-Up must maximize comparability with the content of earlier studies (NLS-72, HS&B) in order to facilitate cross-cohort comparison and analysis of trends.** Just as there is a strong need for content stability across waves of NELS:88, so too is there a need for continuity of content to support trend analyses and comparisons with the NLS-72 and HS&B cohorts.

NELS:88 is the third installment in a continuing program of nationally representative education longitudinal studies. Conceptual, content, and design linkages between NLS-72, HS&B and NELS:88 permit systematic measurement of changes in the secondary education system over time, and facilitate trend comparisons between the experiences and outcomes of the various cohorts. Thus student questionnaires should contain items that are comparable to those on the prior NCES longitudinal surveys of the high school classes of 1972, 1980, and 1982.

Field test staff attempted to identify HS&B and NLS-72 items that have not proved useful to researchers. Such variables are obvious candidates for deletion. Societal changes and shifting policy concerns will render irrelevant some additional items. Nonetheless, a considerable body of measurable and manipulable variables, that have proven analytically powerful in the past and give like promise for the future, remain. These NLS-72 and HS&B items were given high priority for inclusion on the NELS:88 Second Follow-Up questionnaires.

Special steps must be taken when choosing or modifying (or developing) questionnaire items and other measures that are to be used for making trend comparisons. Language and cultural concepts are mutable; twenty years will have passed between the base year of NLS-72 and the 1992 data collection of NELS:88. Sometimes sameness of meaning can be maintained only with some modification in language. Therefore another thrust of questionnaire design in the field test was to examine trend items for stability of meaning and to, cautiously, modify them when necessary, but only to the smallest possible extent compatible with conveying whatever shift of nuance or reference might have occurred over time.

5.) **Key to NELS:88 multilevel design is articulation between data sources.** Three kinds of articulation are required: **first**, some redundancy with contents across Second Follow-Up data sources is desirable, to permit analyses of item validities, perceptual differences, and so on. **Second**, apart from such special cases, redundancy should be minimized, so that, within the limited burden that can be imposed on each respondent population, the maximum amount of information can be obtained. **Third**, cross-wave articulation between the instruments is also necessary -- for example, to obtain information about what happened in the junior year of high school for the student sample, or between 1989 and 1991 for parents (parents were not surveyed in the First Follow-Up), bridging items must be created. These three considerations are elaborated below.

The individual student is the fundamental analytic unit for NELS:88. Student reports are further illuminated by tapping the rich contextual information available from other respondent populations and institutional records sources. Effective analyses of student educational performance and growth require that data from other sources on the school, community, and home environments be included in multivariate models of student status and change. Since the NELS:88 Second Follow-Up will collect data from school records, from students, parents, principals and from teachers, there are multiple sources of information about students' personal characteristics, their backgrounds and environments, and their interactions with their schools.

Such a design offers numerous opportunities for comparing reports from multiple sources. Such "triangulation" is almost always necessary if one is to understand the full impact on key categories of actors of any given school policy, program or initiative, particularly at the level of the perceptions and intentions of affected individuals. Thus, some questions should be asked of principals and of teachers as well as of students. Likewise, a fair measure of deliberate parallelism of items is desirable in the dropout and student instruments to underwrite comparison, while at the same time the divergent experiences of in-school and out-of-school students require that different questions be posed as well. One sort of articulation between data sources that NELS:88 must achieve then is just this sort of desired overlap of items between questionnaires. Another is articulation between questionnaires and cognitive tests. For example, at a class-specific level linked to individual students the NELS:88 teacher questionnaires have inquired into content coverage or "opportunity to learn" ("has this concept been presnted in this class...") in a manner that is linkable to test results.

At the same time, burden to respondents must be limited, yet the NELS:88 design has considerable data requirements, seeking as it does both to engage emergent issues within a changing policy agenda, and to historical continuity with prior waves of the study and with NLS-72 and HS&B. Thus, another major principle of articulation is that, in the main, any given class of respondents should be asked only those questions they can reasonably and, in most cases, **best** be expected to answer. For example, in the Base Year, parents, not students were asked to report family income. An example of this principle directly from the Second Follow-Up is given in the matter of course taking patterns and grades. Since transcripts will be collected, and have been found to be generally a more complete and valid source of curriculum and grades information (Fetters, Stowe and Owings, 1984) than are student self-reports, the Second Follow-Up will not seek student reports of grades and course-taking. In turn, more questionnaire space becomes available for inquiring into additional topics, for which the student is the best information source.

Achieving cross-wave articulation is also essential; there are two issues here. One is that gaps in time need to be bridged. NELS:88 students were most recently surveyed when the majority were in tenth grade. Despite the general avoidance of retrospective questions in longitudinal surveys, information must be obtained -- from school records/transcripts, and from the student -- concerning events in the junior year. The problem of articulation between waves is yet more acute in the parent survey of the Second Follow-Up, because parents were last surveyed in the Base Year.

A second cross-wave articulation issue affects questionnaire design: Which variables need be repeated? Dynamic variables (such intervening or dependent variables as SES, family structure, school environment, aspirations and goals, course-taking, school completion, and so on) must be asked repeatedly. However, certain static variables -- such as age, sex, and race -- are presumed not to change, and these items need to be asked only once. For this reason, all Base Year and First Follow-Up items were re-assessed, to determine their suitability for inclusion on the Second Follow-Up questionnaires.

These five basic study goals and design constraints shaped the development process for all of the NELS:88 field test instruments -- the limited questionnaire time, the need to achieve content breadth in what is fundamentally a general purpose dataset; the need to sustain the primarily longitudinal research objectives of NELS:88, while providing a basis for cross-cohort comparisons to NLS-72 and HS&B; and the need for articulation between the various questionnaires. Issues of policy and process, of content and of form, that relate to each questionnaire individually will be discussed in the separate sections below that pertain to each study component (Section 2.1.1 deals with student questionnaire issues; 4.1 with the dropout questionnaire; 6.1.1 with the parent questionnaire, and 8.1 with the school questionnaire. Note that cognitive test development is discussed in Section 3.10.2.)

While planning and review activities have been described in Section 1.0 and further information on the instrument development process is contained within the account of each individual questionnaire, the briefest of overviews at a more general level, indicating both the instrument design process as it has unfolded thus far and a statement of steps that still remain, may prove helpful.

Essentially, the process of questionnaire and test design in NELS:88 2FU reflects a repeated cycle of seeking recommendations on instrument content and form from substantive experts, policy agencies, and other data users, followed by successive iterations of draft questionnaires. These questionnaire drafts are refined by clinical pretesting and, finally, subjected to a full-scale field test. The field test permits a methodological assessment of question wording and content, formatting and instructions, as prescribed by NCES Standard 87-03-01. While the field test results, as presented in this report, provide a wealth of information about how well the field test items are performing--and numerous suggestions for improving or deleting items--the critical next step in this process is to gather further input from potential data users to ensure that analysis needs guide the choice and inform the final form of items for the main study questionnaires. Hence there is paramount importance in the review of field test results by consultants representing the perspective of educational researchers and policy analysts, by the Technical Review Panel, and by representatives of the various federal and other policy agencies, educational associations and groups, and all other bodies that have an interest in the NELS:88 data.

Nor does the process of methodological refinement cease with this next step of obtaining further substantive input. Further suggestions that arise from the upcoming review process will require further iteration and qualitative assessment and reassessment of questionnaire items. The field test report identifies a number of questions and content areas that require further work. It is our feeling that this work is best performed after further input is received from the Technical Review Panel and NELS:88 consultants. Finally, further analysis and refinement of questions will be called for as the discussion of analytic priorities clarifies which of the many candidates for inclusion on the questionnaire will be chosen. There is a clear benefit to concentrating resources on the improvement of the items that will make the "final cut" rather than investing more heavily in the full pool of items at an earlier stage, knowing that many of these items will not be included in the final questionnaires. In short, completion of the field test report marks the completion of one cycle of substantive and methodological review of the survey forms, but marks as well the beginning of a further such cycle.

### 1.2 Field Staff Recruiting and Training

Recruiting and hiring the field staff for the NELS:88 Second Follow-Up Field Test was conducted by NORC's Office of Field Coordination and Management NELS:88 Liaison, in conjunction with the field test task leader. Field personnel were recruited for two positions: 1.) field managers who were responsible for district contacting and supervising the work of the field interviewers, and 2.) field interviewers (or team leaders) who contacted schools, visited each school during the fall, and conducted in-school and off-site data collection sessions. For both positions, priority in hiring was given to staff who had performed satisfactorily on the NELS:88 Base Year and/or First Follow-Up. In the instances that this level of experience could not be achieved, individuals were hired who were experienced in group administration or other data collection for NORC.

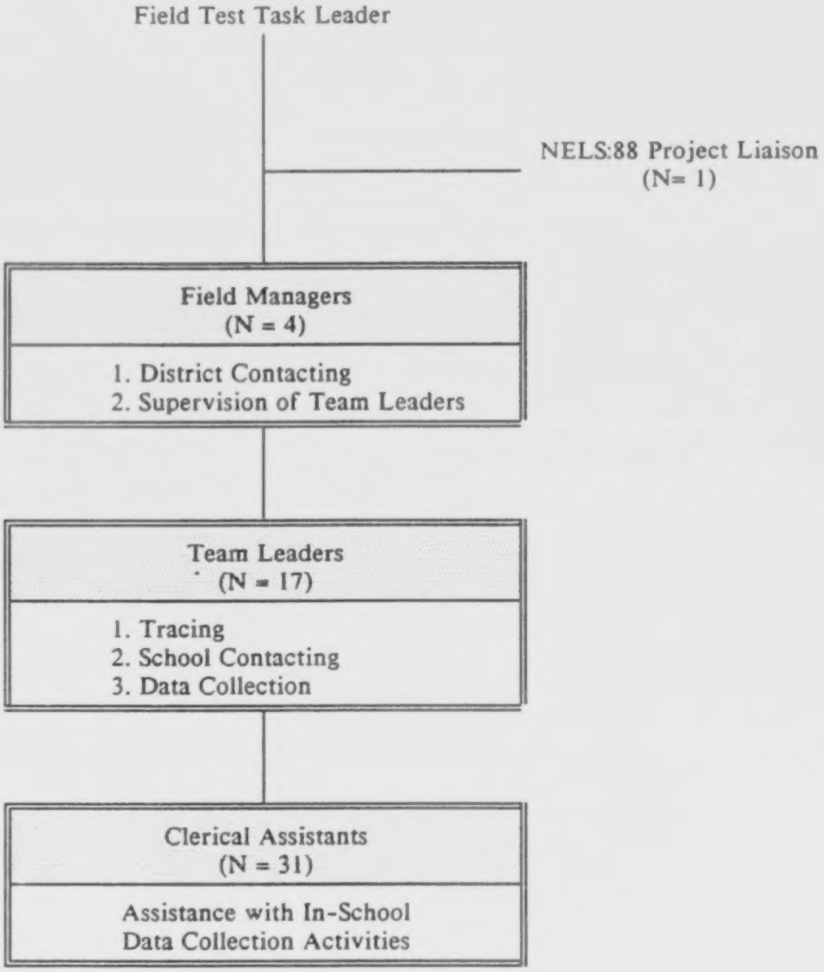
Staffing of the project field component began with the hiring of four field managers (FMs). Three of the Field Managers were assigned field test activities for an entire state (California, Texas, and Florida); one Field Manager was responsible for two field states (Illinois and New York). Of these four field managers, three had worked on the NELS:88 Base Year and First Follow-Up. The non-NELS experienced field manager had had experience in the same role on HS&B. As noted above, these field managers contacted district superintendents and supervised school contacting, tracing, and survey administration activities.

The recruiting and hiring of field interviewers was also conducted by NORC's Office of Field Management and Coordination NELS:88 Liaison and Divisional Field Managers, in conjunction with the field test task leader. Recruiting instructions specified that as many interviewers as practical should reside in the communities containing the 1FU field test schools -- schools likely to be asked to participate in the 2FU field test. It was also specified that these interviewers would be assigned both the fall school contacting (securing cooperation and sample freshening and augmentation) and winter data collection activities, assuming their performance in the fall was satisfactory. Previous NELS:88 experiences have suggested that this continuity of assignment promotes mutual trust and smoother working relationships between the interviewer and the school staff, thereby facilitating sample freshening, tracing of cohort members, and acquisition of other student file information. Debriefings conducted with interviewers confirm that this continuity of staff still facilitates project work. (See Figure 1.2.1 for further explanation of the Second Follow-Up Field Test staffing structure.)

Seventeen interviewers were hired to work on the Second Follow-Up Field Test. These were distributed among the five field test states as follows: seven in California, two in Florida, three in Illinois, two in New York, and three in Texas. Of these 17 interviewers, 15 had had experience on prior waves of NELS:88; the remaining two were experienced NORC interviewers.

During preparations for the in-school data collection (Survey Day) sessions, field interviewers were directed to hire clerical assistants for help in administering the student survey. These clerical assistants were responsible for giving general clerical support to the field interviewer. These support activities included: passing out questionnaire and test booklets to the students, distributing

**Figure 1.2.1 NELS:88 Second Follow-Up Field Test Staff Structure**


Field Test Task Leader
NELS:88 Project Liaison
(N= 1)
Field Managers
(N = 4)
1. District Contacting
2. Supervision of Team Leaders
Team Leaders
(N = 17)
1. Tracing
2. School Contacting
3. Data Collection
Clerical Assistants
(N = 31)
Assistance with In-School
Data Collection Activities

refreshments during the break, and completing the critical item edit during the testing session. Discussion of selection criteria and performance expectations was a part of the interviewers' training review with their field managers.

Materials for self-study training were developed for each specific activity (e.g., District Contacting Manual, Tracing Manual, School Contacting Manual, Team Leader Data Collection Manual, and Clerical Assistant Manual). These training materials were developed jointly by the central office project staff and the field managers. NORC has employed the combination of self-study and telephone review with interviewers on many studies; these training techniques have proved reliable and inexpensive.

The manuals included discussions of all aspects of the study, from the broad research objectives of NELS:88 to task-specific procedures, and of the role and expectations of the field staff. Self-study exercises were enclosed with the manuals; these self-tests were designed for interviewers to assess their own knowledge of the information in the manual. Once interviewers had carefully read the manual and had complete the self-study exercise, a telephone review was scheduled with their field manager. The field managers reviewed procedures, answered questions, clarified problematic points, and assessed the interviewer's understanding of the procedures. These training assessment questions dealt with issues of procedure and with challenging situations that might arise in the schools.

### 1.3 Sample Design and Selection

The NELS:88 Second Follow-Up (NELS:88 2FU) field test sample was designed to satisfy two primary objectives: first, to obtain a sufficient number of observations to assess the survey instruments and field procedures and, second, to augment the field test cohort with enough additional twelfth graders to accurately assess the free response cognitive test items.

In order to meet these two objectives, it was necessary to construct the sample in three steps. In the first step, all NELS:88 field test cohort members (those selected into the Base Year Field Test initial sample and individuals added to the cohort through sample freshening in the First Follow-Up field test) were identified for the NELS:88 2FU field test. Second, individuals were added to the sample via the 2FU freshening procedure performed at the twelfth grade schools attended by field test cohort members. Third, in order to have enough free response item observations, additional individuals were randomly sampled from school rosters at NELS:88 2FU field test schools.

The following paragraphs describe the sampling procedures that were used to identify the NELS:88 2FU field test sample members.

#### 1.3.1 NELS:88 Field Test Cohort

All NELS:88 field test cohort members selected during the Base Year and individuals added to the field test cohort via freshening during the First Follow-Up field test were retained for the Second Follow-Up field test sample. The total number of NELS:88 First Follow-Up field test cohort members was 1,086. Of this number, 1,031 were cohort members selected for the Base Year field test and 55 were added to the cohort through freshening during the First Follow-Up field test. Without the addition of free response cognitive test items to the field test design, retaining all cohort members and freshening during the 2FU field test would have provided an adequate number of cases for assessing the student questionnaire, field procedures, and multiple choice cognitive items.

### 1.3.2 Sample Freshening

The purpose of the sample freshening procedure in the main study is to make the NELS:88 2FU sample representative of the twelfth grade class of 1992. The freshening procedure was field tested by sampling current twelfth grade students who were not eligible for inclusion in the eighth grade field test sample in 1987 because they were not in the eighth grade or not in the U.S., in 1986-87. Following is a detailed description of the freshening procedure and a summary of the results.

**Stage One Freshening:** The procedures used to freshen the twelfth grade field test sample borrowed heavily from 1FU sample freshening procedures, with an additional step to ensure high quality data regarding the eighth grade "anchor point".

In the first step, a roster of the names of all twelfth grade students was obtained by an NORC representative during her or his fall visit to the 2FU field test school. (When necessary, an alphabetical list of all twelfth graders was then constructed.) The NORC interviewer then found the name of the first sample member on the roster and inquired of school staff if the next student listed after the sample member was in the tenth grade in the U.S. in 1989. If a student not in the tenth grade in the U.S. in 1989 was found, he or she was added to the sample for data collection. The student listed directly under this candidate for stage-two freshening was checked using the same procedure until the link was broken (a student was found to have been in the tenth grade in the U.S. in 1989). This procedure was circularized so that when an interviewer reached the last student on the roster, she or he returned to the first student on the roster and continued the process until reaching the first NELS:88 sample member. The procedure was repeated for each sample member on the roster.

**Stage Two Freshening:** The original sample members plus the stage-one freshened students constituted the 2FU field test cohort from which data collection was attempted. An additional step was implemented in response to concerns that school staff would not know with accuracy the status of a student after four years and a transition to another school. In the second step, response to two questions (Q14a. and Q14b.) in the New Student Supplement (completed by all newly added participants) was checked. This item asked the student if he or she had been in the eighth grade (subpart a.) and in the U.S. (subpart b.) during the spring term of the 1986-87 school year. Those students who answered "no" to **either** question were then permanently added to the field test cohort as freshened students; the data from those responding "yes" to both items were discarded. When this freshening procedure is implemented in the main study, the selection probability for the freshened student will be equal to the selection probability for the original NELS:88 sample member that preceded him or her on the roster.

**Results:** The stage one sample freshening procedure implemented in the Field Test added a total of fifteen students to the data collection effort. The fifteen stage-one freshened students contributed less than two percent of the cohort data collection, although there was some variation among the five states (CA 1.2%, FL 2.6%, IL 0.8%, NY 2.9%, and TX 0.5%).

Six of the 15 stage-one freshened students participated in a Survey Day or Make Up Day. Of these respondents, two answered Questions 14 a. and b. indicating that they had not been in the eighth grade in the spring term of the 1986-87 school year or had not been in the U.S. These two students were permanently added to the field test cohort. Of the six, it was determined that information about the tenth grade status of the student had been inaccurate in two instances, leading to further doubt that high schools will be capable of providing accurate eighth grade status data on which to base freshening of the 12th grade main study cohort. Consequently, we will not rely soley on 12th grade school data for sample freshening.

However, it is doubtful that the main study freshening procedures will repeat the two step freshening process exactly as used in the field test. Rather than collecting all Survey Day data from stage-one freshened students, we are considering a very brief contact with these students by telephone or mail to determine their status in the 1986-87 school year and hence their eligibility for freshening.

### 1.3.3 Sample Augmentation

One of the major objectives of the 2FU field test was to determine the appropriateness and examine the psychometric properties of free response cognitive test items for twelfth grade students. As mentioned above, the design of the field test cognitive test battery was such that the 1,086 member field test cohort did not provide the number of cases to analyze these items. Hence, it was necessary to augment the field test sample to obtain the necessary 2,000 observations. To ensure a total of 2,000 completions within budgetary and schedule constraints, the sample size was increased to approximately 3,600. The augmented 2,500 students were temporary additions to the sample for data collection, unlike the two freshened students who will become permanent members of the longitudinal cohort. The following paragraphs describe the sample augmentation procedures.

For each of the 105 schools that agreed to participate in the 2FU field test, NORC interviewers sought permission to randomly select additional students to bring the total number of students selected for the 2FU field test to 40. Sample augmentation was refused by a total of nine schools; two other schools were not included in the field test because of the small number of students attending the twelfth grade.

As with sample freshening, sample augmentation procedures utilized school rosters. In preparation for augmentation, the interviewer first crossed off the roster the names of NELS:88 cohort members (including candidates from stage-one sample freshening) and of students who were ineligible to complete the survey. (This last criterion was eliminated after trials in several schools demonstrated that school staff could not easily identify ineligible students. The procedure was then changed so the eligibility check was performed after the augmentation selection.) After the NELS:88 sample members had been crossed off, the interviewer numbered the roster from 1 to n (with n being the last student on the roster), assigning a number to each eligible non-NELS:88 student.

After the roster was numbered, the interviewer then chose the augmented students, utilizing sample selection tables generated specifically for each school. The sample selection tables, produced by NORC's senior statistical scientist, contained the school's calculated grade size and 60 random numbers to be used for sample augmentation. The calculated grade size was an estimate of the school's senior class total enrollment, obtained by dividing the total school enrollment (as reported by Quality Education Data) by the number of grades, thus assuming that the senior class enrollment would be approximately one quarter of the total school enrollment in a school with four grades. The tables were constructed such that the random numbers between 1 and the calculated grade size were arrayed in random order.

Next, the interviewer identified the total number of non-NELS:88 students available for augmentation and determined the number of students to augment by subtracting the number of cohort members from the target cluster size of 40. The NORC representative then systematically went through the list of random numbers, moving from left to right and top to bottom, assessing each number's applicability (e.g. - if the random number was less than the total number of eligible, non-NELS:88 12th graders, they circled that number on the selection table; if the number was larger, they marked a line through it.) The interviewer continued through the table until enough numbers had been circled to augment the sample to a total of 40 students. For each random number circled, the interviewer identified the student on the roster who corresponded to the number. For each student selected, the interviewer recorded the student's name from the school roster and collected his/her parents' names, address, and telephone number from the school files.

Given that 2FU field test goals differed from those of the 1FU field test and that there are the natural limits on school cluster size, the number of schools needed to participate in the 2FU field test was higher than the 75 schools which hosted Survey Day sessions in the 1FU field test. To obtain the number of cognitive test observations needed, the 2FU field test set a goal of 95 - 100 participating schools. It was apparent that it would be necessary to seek the cooperation of schools that had not participated in the First Follow-Up field test. These were primarily schools to which one or two field test sample members had transferred and the transfers were the only NELS:88 cohort members at the school. Because augmentation tables were only produced for schools that participated in the First Follow-Up field test, it was necessary to produce a second set of augmentation tables for interviewers to use in augmenting the student sample at these schools new to NELS:88.

The second set of augmentation tables was produced using procedures identical to those employed in the generation of the first set of tables. An SPSS PC program created a table of 60 random numbers for each new school based on the schools' twelfth grade enrollment determined by interviewer's contact with school staff. The tables were similar in appearance to those in the first set. Interviewers used the tables in the second set in the exact same way as those in the first set were used.

It was also necessary to supplement some interviewers' tables when the original tables did not provide enough numbers to perform sample augmentation. When this happened, numbers were provided to interviewers from the table of random numbers (Table B) in Blalock (1979). The numbers were chosen by randomly choosing a start column and then selecting all of the numbers in that column. One column generally contained sufficient numbers to adequately supplement the interviewers' tables. All additional numbers given were recorded by the interviewer on his or her copy of the table as well as on the central office copy of the original augmentation table.

### 1.4 Student Tracing

Because NELS:88 is a longitudinal study, the successful tracing of sample members during each wave of the study is a critical activity. Thus, successfully locating a large percentage of the sample members and thoroughly documenting their whereabouts in the 2FU will help maintain the integrity of the longitudinal sample and will significantly facilitate the task of tracing sample members in the Third Follow-Up. In the fall of 1990 (September 25 - December 14), NORC field staff attempted to locate each member of the NELS:88 field test cohort. The first step of tracing was to determine whether or not sample members were enrolled in school. For those who were found in school, the primary aim of tracing then became documenting in which school sample members were enrolled. For those not in high school, information was collected that documented the sample member's status as a dropout or early graduate. Locating information was verified for all traced sample members.

Students were traced using three different treatments or methods: batch, single, and 1FU nonrespondent. In the First Follow-Up field test student questionnaire, students were asked to name the school they were most likely to be attending in the twelfth grade. The school they named was referred to as the "nominated school". Students assigned to the batch treatment were grouped together by their nominated school and traced through the school as a group.

Students who did not expect to attend the same school as most of their tenth grade classmates were assigned to the single treatment. These students were traced individually; the first contact for tracing was made through the household. Students initially assigned to the batch treatment who were not found at their nominated school became single treatment cases.

The third treatment used for tracing involved 1FU nonrespondents, also referred to as NRs. These sample members were in the 1FU field test sample but did not participate. NR cases were treated in much the same manner as single cases, the major difference being that tracing efforts began at the 8th or 10th grade school rather than the household, since household information was not available for the majority of these cases.

Field interviewers were given an information sheet for each case containing locating information that was provided by sample members in the First Follow-Up field test. Specifically, these sheets included: the sample member's full name, race, sex, date of birth, nickname, home address, and telephone number; the names, addresses, and telephone numbers of both parents/guardians; and the name, address, telephone number and the relationship of the person the sample member identified in the First Follow-Up field test questionnaire as being their "closest relative or friend". In addition to the locating information, the sheets also contained the names, addresses, and phone numbers of the sample member's 8th and 10th grade schools; the name, address, telephone numbers, and ID numbers of the sample members' nominated twelfth grade school; and the name of the school that was nominated most often by the sample member's tenth grade classmates. Finally, the sheet also contained information about the sample members' participation in the Base Year and First Follow-Up field tests.

Field interviewers kept detailed accounts of all contacts with sample members, household members, and school personnel in order to maximize tracing efficiency and to facilitate the communication of case details to their supervisor for technical assistance or transfer of the case to another field interviewer. Each interviewer participated in weekly conferences, in which aggregate production figures, as well as the status of each case, were reported to his or her supervisor.

After a student was successfully located, field interviewers recorded student and school identification numbers and all locating information on a tracing form. These forms were sent to the central office and entered into the 2FU field test Survey Management System (SMS) data base.

### 1.4.1 Tracing Procedures

**Batch Treatment:** Interviewers were instructed to give their batch treatment cases top priority, because tracing at the schools would identify sample members who needed to be traced individually. The batch treatment involved calling the nominated high schools of field test sample members to ascertain their enrollment status and to verify the school name, address, and principal title and name.

Interviewers' first contact in batch tracing was with the schools' main office administrative assistant. A script was provided to interviewers that guided them through their interactions with school personnel. After identifying themselves as NORC staff persons calling on behalf of the U.S. Department of Education, interviewers explained that they were seeking assistance in determining whether the NELS:88 sample members were enrolled in the school. Interviewers stressed that they were not collecting data from respondents at that time. (If the administrative assistant indicated that it would be necessary to obtain the approval of the principal to perform tracing, the interviewer attempted to set an appointment with the principal to seek such approval.)

Once schools agreed to tracing, interviewers read the names of the NELS:88 students thought to be in the school to a school staff member. When necessary, interviewers used data from the information sheet to confirm a student's identification. For each student, interviewers recorded whether the student was enrolled or not enrolled in the school. Finally, interviewers verified the correct school name, address, and principal title and name. They also confirmed the school mailing address. When school principals could not be persuaded to cooperate with tracing efforts, all students listed at that school were reassigned to the single tracing treatment.

**Single Treatment:** Students who nominated a school other than that nominated most often by their tenth grade classmates were assigned to the single tracing treatment. Interviewers first attempted to contact the sample member or his or her parent or guardian by calling the household. If contact was made, interviewers attempted to confirm the sample member's enrollment status, the name of the school at which the sample member was enrolled; and the sample member's address and telephone number. If the interviewer could not reach the sample member or his or her parent or guardian, an attempt was made to get better locating information from a close friend or relative of the sample member.

After all attempts to trace the student through the household failed, interviewers attempted to trace the student through schools, beginning with the students' nominated school and, as tracing efforts failed, continuing with their tenth grade school, eighth grade school, and the modal school of their tenth grade classmates.

If enrollment status at a school was obtained from a source other than the school itself, the school's location, mailing address, and principal title and name were verified through a brief telephone call to the school. Field interviewers recorded all final tracing information obtained on forms that were mailed into the central office on a weekly basis.

**Nonresponse Cases:** Sample members assigned to the NR tracing treatment included 1FU field test nonparticipants and 1FU field test participants who did not nominate a twelfth grade school. First Follow-Up field test nonparticipants included those who missed survey sessions and those who refused to take part in the study. Interviewers were instructed to be sensitive to the special concerns of sample members who refused or whose parents refused to participate in the 1FU field test.

The first point of contact for tracing NR cases varied depending on whether the sample member was a 1FU field test nonparticipant or was a participant who did not nominate a twelfth grade school. For 1FU pretest nonparticipants, the first point of contact was the household, and tracing progressed in the same manner as single treatment students. For 1FU participators who did not nominate a twelfth grade school, the first point of contact was with the school that was nominated most often by the sample members' tenth grade classmates. Tracing procedures for these students were the same as those used for sample members assigned to the batch treatment. Again, school data obtained from any source other than the school itself were verified by calling the school. All final tracing information was recorded on tracing forms and mailed to the central office on a weekly basis where it was entered into the field test SMS.

### 1.4.2 Student Dispersion

One of the major challenges of any longitudinal study is to maintain contact with cohort members. Typically, the primary difficulty is to locate cohort members who have moved since the study was last in contact with them. An assessment of the nature and degree of dispersion of the field test cohort that occurred between the tenth and twelfth grade was one of the major goals of the field test tracing effort. It seemed that this assessment would provide both qualitative and quantitative information that could be used to plan the 2FU main study tracing effort.

34

The tracing effort found most (87.5 percent) of the field test cohort members in high school. Of those who were in school, eighty-six percent were at their tenth grade school; three percent had transferred to a school that contained other field test cohort members; and eleven percent had transferred to a school that did not contain any other field test cohort members. The remaining cohort members fell into four categories: dropouts, those who were at home incapacitated (one cohort member), those who had moved out of the five field test states or out of the country entirely, and those whom NORC could not locate. A total of 68 cohort members (six percent of the total cohort) were found to be dropouts (see section 1.7 regarding the identification of dropouts) during the fall tracing phase. NORC successfully located all but 14 of the 68 dropouts. Fifteen cohort members either moved out of the five field test states or out of the country. Finally, approximately five percent of the field test cohort (59 individuals) could not be located; of these 59 cases, seven were never fielded for tracing due to a dearth of information (for example, records from previous waves included only a name and ID number).

### 1.4.3 Tracing Results

The primary goal of tracing was to determine which high school each sample member was attending or to confirm a sample member's dropout status if they were not found to be attending school. Overall, ninety-five percent of all cohort members were successfully traced. As suspected from our experience in the 1FU, successfully tracing freshened students was more difficult than for the core sample membership; while successfully locating 95.8 percent of core sample members, the rate for students added in the 1FU through freshening procedures was 78 percent. We believe that this is a combined effect of several factors: 1.) freshened students are more likely to move than core students; 2.) freshened students are more likely to drop out of school than core students; and 3.) because freshened students were less likely to participate in the 1FU, we have less data about these students to use in efforts to locate them. Field interviewers were asked to report the point of contact through which each sample member was traced. This information was recorded on tracing forms and sent to the central office. Table 1.4.1 below summarizes the information sent in by the interviewers.

**Table 1.4.1: Cases* Successfully Traced by Initial Point of Tracing Contact**

| Point of contact | Frequency | Percent |
|---|---|---|
| Nominated school | 768 | 76.4 |
| Student or household | 118 | 11.7 |
| Other school contact | 102 | 10.2 |
| Other type of contact | 17 | 1.7 |

*Note: Due to interviewer error, this information was not provided for 20 of the cases successfully traced.

Source: NELS:88 Second Follow-Up Field Test Data.

Not surprisingly, the table illustrates that most students (76 percent) were found through contact with the school they nominated in the 1FU field test. The point of contact second most frequently cited was the student or a member of the student's household. It was necessary to trace sample members through the household in nearly 12 percent of cases. Ten percent of the students were found through contact with a school other than their nominated school. For the most part, these were students who were not attending the same school they attended in the tenth grade. Finally, nearly two percent of all cohort members were traced through some "other" contact person. "Other" contacts included departments of motor vehicles and relatives not living in sample members' households.

### 1.5 Securing State, District, and School Cooperation

Securing permission at state, district/diocese, and school levels to conduct the NELS:88 2FU field test in sampled schools was a significant task of the fall 1990 activities. Each school sampled can be categorized into one of three sector types, each of which has its own organizational structure: public schools, Catholic schools, and non-Catholic private schools. Differences in organizational structure and levels of authority required that the task of securing cooperation be tailored to each sector type.

For public schools, it was first necessary to secure the cooperation of the Chief State School Officer of the state education department, followed by the district superintendent, on occasion a subdistrict, and finally the principal of the sampled school. Since each level in this hierarchy exercises some degree of autonomy from the next higher level, securing permission at one level to conduct the study did not guarantee cooperation at subordinate levels.

Catholic schools exist within a more limited hierarchy of authority consisting of only the diocese, archdiocese, parish, or other religious group and the individual school. Consequently, it was only necessary to secure cooperation at these two levels. Finally, non-Catholic private schools seldom acknowledge any higher authority; in rare cases they may be subordinate to a local authority, such as a religious organization. Securing cooperation began at the school level for schools of this sector type.

#### 1.5.1 Procedures for Securing Cooperation

**Securing Endorsements from Professional Organizations:** Prior to approaching state departments of education, a letter of endorsement was sought from the Education Information Advisory Council (EIAC) of the Council of Chief State School Officers. Dr. John Stiglmeier of the New York State Department of Education and Dr. Gordon Ensign of the Washington State Department of Education reviewed the field test design and instruments on behalf of EIAC. Mention of approval of the survey by EIAC was incorporated into the state contacting letter sent to Chief State School Officers. Endorsements were also sought and received from: the American Association of School Administrators, the National Association of School Boards, and the National Association of Secondary School Principals.

Endorsement of the study by the National Catholic Education Association was also sought prior to the contacting of dioceses and archdioceses. Mention of the association's approval was made in the diocese contacting letters and in the initial contacting call to Catholic school principals.

**Securing State Cooperation:** In August of 1990, the NORC Project Director sent letters to the Chief State School Officers (CSSO's) in the five field test states--California, Florida, New York, Texas, and Illinois--requesting each state's cooperation in the study. The letter and accompanying informational materials explained the study's design and purpose, and mentioned the approval of the field test by the EIAC. The letter requested that the CSSO appoint a state coordinator for the field test. The state coordinator was responsible for keeping the state informed of the progress of the study, handling inquiries from district superintendents concerning state approval, and advising project staff regarding survey-related problems encountered at the district and school levels. Permission to contact districts was easily secured from CSSO's in every state, and all CSSO's appointed a state coordinator.

**Securing Public School District Cooperation:** District cooperation activities began in September 1990 with the mailing of introductory letters to district superintendents; these letters requested the superintendent's permission to contact the principals of selected schools. (In the letters to districts which had participated in the First Follow-Up Field Test, superintendents were reminded of their participation in that study and were asked for their continued cooperation.)

One of the four NORC field managers followed up on the letter with a telephone call to each superintendent. The field managers used as their guide the District Cooperation Manual containing contacting scripts, recommendations for circumventing gatekeepers, refusal aversion/conversion advice, and copies of the letters sent to CSSO's and superintendents.

During a telephone contact with the superintendent, the field manager requested the superintendent's permission to contact the principals of selected schools in the district. During this call, the field manager mentioned the need to freshen and augment the NELS:88 core sample. Upon securing the district's cooperation, the field manager asked the superintendent to appoint a district coordinator. The responsibilities of the district coordinator, the field manager explained, included handling principals' inquiries about district approval.

When a superintendent refused to approve contacting the principal, the field manager attempted to address the superintendent's concerns about the study. Persistent refusals were referred to upper-level project staff for conversion, and also to the state coordinator.

Once field managers had secured district cooperation, the corresponding schools were released to field interviewers, supervised by the four field managers, for securing cooperation at that level.

**Securing Diocesean Cooperation:** Procedures for securing diocese approval differed little from those implemented for securing district cooperation. Each diocese containing a selected school was sent a letter of notification about the study, and follow-up calls were made by one of the four field managers. However, diocese officials were not asked to appoint a coordinator, because the project perceived no need for one, given the autonomy displayed by a majority of Catholic schools.

**Securing School Cooperation:** Field interviewers contacted the principals of sampled public and Catholic schools shortly after the district or diocese agreed to cooperate. Non-Catholic private schools were immediately released to field interviewers for securing cooperation, since permission to contact headmasters and headmistresses was not required from a higher level.

During the initial telephone contact with the principal, the field interviewer explained the nature and purpose of the Second Follow-Up Field Test, particularly the necessity of freshening and augmenting the student sample. Principals were told that, with freshening and augmentation, the number of students to be surveyed would be 35 to 40. Interviewers addressed principals' questions and concerns in this call and requested their permission to conduct the study in the school during February or March of 1991. Field interviewers also attempted to avert and convert refusals.

Upon agreeing to participate in the field test, the principal was asked to appoint a school coordinator (whose duties the interviewer briefly outlined) and to schedule a Survey Day, on which the student survey would be administered, and a Make-Up Day, on which students absent on Survey Day would be surveyed. The field interviewer also scheduled a visit to the school with the principal, explaining that during the visit the interviewer would procure a course catalog, verify the enrollment of the student sample, and freshen and augment the sample. The interviewer asked the principal either to mail him or her a copy of the twelfth grade student roster or to have a copy available at the time of the visit.

One interesting finding of the 2FU Field Test relates to the scheduling of Survey Days. One might predict that Survey Days would be randomly distributed over the school days within the data collection period (February 4 - March 23). However, most school principals chose school days in the early weeks of the field period, citing other tests scheduled, other senior class activities, and the mindset of seniors as they approach graduation. We anticipate a similar scheduling phenomenon in the 2FU Main Study.

After speaking to the school principal, the interviewer called the school coordinator and explained his or her duties: to assist with sample freshening and augmentation during the school visit; verify permission type and track permission forms; distribute Survey Day reminders to students; receive, check, and store Survey Day materials; and attend the Survey Day session.

**Field Procedures and Budget Constraints:** During the securing of school cooperation, several procedural modifications were implemented to maximize the number of NELS:88 and augmentation students surveyed while containing costs. These are described below.

1.) When, during tracing, an interviewer determined that only one NELS:88 student attended a school, district and school cooperation were, as usual, sought by telephone; however, the interviewer did not schedule a school visit, nor ask the principal to set a Survey Day or appoint a School Coordinator. These cases were put on hold until later subsampling of single student schools.

2.) When a school principal refused to allow the interviewer to freshen and augment the sample, and the school contained 25 or more NELS:88 students, the interviewer was instructed to drop augmentation and to offer only to freshen the sample. If the principal still refused freshening, the interviewer documented the principal's objections to freshening. If a school refusing freshening and augmentation contained fewer than 25 NELS:88 students, the interviewer informed the principal that it would not be cost effective to send an interviewer to the school to conduct Survey Day, and attempted to arrange a school coordinator-administered Survey Day.

3.) Interviewers also requested that school coordinators conduct Survey Days when the twelfth grade class was small and the sample could not be augmented to 30 students.

4.) Several schools that were added to the sample during the subsampling of single student schools were eliminated from the field test because of their distance from field test interviewers. Interviewers were instructed not to pursue cooperation from schools more than two hours from their homes.

5.) The number of Make Up Days conducted by field interviewers was limited to contain data collection costs. Each of the four field managers was permitted a maximum of three interviewer-administered Make Up Days, to be disbursed to interviewers as deemed necessary. Field managers were instructed to allocate these interviewer-administered Make Up Days in such a way as to maximize the number of students surveyed.

### 1.5.3 Results of Securing Cooperation Activities

All five states (California, Florida, Illinois, New York, and Texas) agreed to cooperate in the Second Follow-Up field test. Of the 63 districts contacted, five refused to allow their school(s) to participate in the NELS:88 2FU field Test. This represents a 92 percent district participation rate.

Of the 108 schools contacted, 95 agreed to host a field test Survey Day session. Seven refused to participate. One school imposed special conditions that NORC was disinclined to accept. Six additional schools did not participate, although willing to do so, for various reasons including: too few 12th graders to be a cost effective contribution to the field test and an unwillingness to allow sample freshening or augmentation. An 88 percent participation rate was achieved by the close of the school cooperation activities in December 1990. One additional school refused during data collection, for a total of 94 participating high schools.

The results of field test efforts to secure the cooperation of districts and schools are not nearly as applicable to the Second Follow-Up main study as seeking state cooperation and other field test activities. The primary difference that makes generalization precarious is the necessity to augment the sample to increase cognitive test observations. (See Section 1.3.3 for a description of sample augmentation.) Second Follow-Up main study schools will not be asked to add students, except through sample freshening which increases school clusters by only one student per school on average. Therefore, we anticipate that district and school cooperation rates will be higher in the main study than those achieved in the field test.

**Special Conditions Requested by Schools:** Several schools required some variation in procedures as a condition of their participation. These special conditions included: examination of the questionnaires and permission forms (2), January (or later) selection of a Survey Day (2), prohibition on release of parent address information (5), interviewer distribution and receipt of permission forms (4), and a snack specified by the school (1). NORC complied with these requests. One school, with a cluster of 30 field test sample members, imposed conditions for their participation that NORC was unwilling to accept; the school requested that the Survey Day be scheduled for a Saturday and that each student be paid 15 dollars for participation.

## 1.6 Obtaining Parental Permission: Implicit and Explicit Consent

**In-School Sessions:** During the initial contacting call, field interviewers asked principals which type of parental permission the school would require: implied permission, which required parents to return a form only if the parents did not want their child to participate; or explicit permission, which required all parents to return the form, indicating whether or not they approved of their child's participation in the study. If there was no state, district, or school policy requiring explicit permission, the use of implied permission forms was strongly recommended. The field interviewer explained to the principal that the implied permission form does not compromise the rights of parents to object to their child's participation in the study, and that our past experience has shown this type of permission to be less burdensome to school personnel. Tracking unreturned explicit consent forms, it was stressed, could be expensive and time-consuming. The interviewer also pointed out that the explicit form would adversely affect the overall participation rate of student respondents and consequently compromise the quality of the data collected. Four of the 94 participating schools required explicit parental permission; principals of the remaining 90 schools chose to use implicit permission forms.

**Form Distribution/Tracking Procedures:** Student packets containing permission forms were sent to the school coordinator two weeks prior to Survey Day. The coordinator distributed these packets to the selected students and instructed them to take the forms home to their parents. Permission forms were to be returned to the school coordinator, who kept track of returned forms. In the case of schools requiring explicit permission, the school coordinator was asked to follow up on forms that had not been returned and to send another permission form home with the student. Team leaders checked permission status on Survey Day and mailed collected forms (or copies) to NORC.

**Parent Conversions:** Team leaders called school coordinators three days prior to Survey Day to determine which parents had refused to permit their children to participate. Team leaders then called these parents (student telephone numbers had been collected during the initial school visit) to convert them. Converted parents were sent an explicit permission form to verify the conversion.

**Off-Campus Sessions:** A special explicit parental permission form was mailed to respondents who were to be surveyed off campus (dropouts), as part of an introductory packet. Parents were instructed to return the signed form to the team leader by mail prior to the day of the off-campus session, or to send the completed form with their child to the session.

### 1.7 Identification of Dropouts and Early Graduates

Two populations of particular interest in the NELS:88 Second Follow-Up are high school dropouts and early graduates (students who take fewer than the prescribed number of years to graduate from high school.) Steps were taken in the 2FU field test to develop and test procedures to identify dropouts and early graduates.

**Dropouts:** Field staff identified dropouts based on information they obtained in their contacts with schools and household members. They were given the following 1FU definition of a dropout to use in classifying 2FU field test sample members.

**Definition of a Dropout: A student is considered a dropout if he or she has not attended school for the last (consecutive) 20 school days (excluding any excused absence).**

When a school official identified a sample member as a dropout, interviewers were instructed to contact the household to confirm the status of the sample member. If an adult household member indicated that the definition above was applicable, the sample member was classified as a dropout. Similarly, if sample members themselves told field interviewers that they were dropouts, they were classified as dropouts. This policy of confirming status through the household was instituted in previous tracing activities, as we found that school-provided dropout status was often in error.

When a dropout was identified according to the procedures outlined above, the interviewer obtained the date the student dropped out of school during the contact with the parent/guardian or sample member. This information was sent to the NORC office on a tracing form and entered into the Survey Management System. A total of 68 dropouts were identified during fall tracing activities. Of these, 54 were confirmed through the household with a parent/guardian or the sample member, and 14 were not located.

**Early Graduates:** Unlike plans for the Second Follow-Up main study, a supplement to the Student Questionnaire was not developed for early graduates. Even though no data were collected from these individuals, it was important to develop and test procedures to identify early graduates in the field test. NELS:88 field test cohort members were in the eighth grade in 1987 and would be expected to be in the twelfth grade in the fall of 1990 if they followed the typical rate of progression through high school. During fall tracing activities, several school officials identified NELS:88 sample members who fell outside of this traditional high school sequence. As interviewers encountered out of sequence students, they recorded the nature of the situation (held back or early graduate) on a tracing form which was forwarded to the central office. The information about out of sequence students was subsequently entered into the field test SMS.

This process undertaken in the fall tracing activities identified a total of 14 early graduates. As with dropouts, an additional number of early graduates were discovered during interviewers' pre-Survey Day calls with school coordinators. During these calls an additional 25 early graduates were identified. In addition to early graduates, school coordinators also identified a total of 26 students out of sequence -- four students in the tenth grade and 22 students in the eleventh grade.

## Chapter 2: Student Survey Procedures

This chapter describes and evaluates the data collection and preparation procedures for the student component of the Second Follow-Up Field Test. Student survey administrations -- primarily conducted by NORC field staff -- were held in selected schools in the five field test states from February through March of 1991.

The objectives and content of the Student Questionnaire are initially reviewed. Following a brief discussion of the methodological experiments integrated into the questionnaire, data collection procedures are elaborated. The success of these procedures is assessed on the basis of field staff reports and project staffs' evaluations. Specific recommendations are made for changes in these procedures, to reduce burden to school staff, to increase student participation in the study, and overall to improve the quality of the data collected.

### 2.1 Objectives, Content and Experiments in the Second Follow-Up Field Test Student Questionnaire

**Objectives and Contents:** Since the student is the basic analytic unit in NELS:88, the student questionnaire, tests, and individual records sources (such as academic transcripts) are the cornerstone documents of the study. Other questionnaires and data sources -- teacher, parent, school administrator questionnaires; archival data on school course offerings and enrollments -- operate as extensions of the student instruments; their primary purpose is to gather information on various school and teacher (or teaching) characteristics likely to influence student learning and persistence in school.

Specific questionnaire content for NELS:88 has been shaped by the study's research agenda. That agenda encompasses institutional and individual-level issues and characteristics--basic research issues and policy and program evaluation issues. It embraces both descriptive statistical needs, and the data requirements of analytical models, particularly those that seek to elucidate educational processes and to identify and explain cause-and-effect relationships between schooling and student outcomes. While hardly an exhaustive list, perhaps the most important broad issues to be addressed by the study are these seven:

1. Equity/Access/Choice;

2. Cognitive Growth and its Correlates;

3. Ability Grouping/Tracking;

4. Dropout Phenomena: The dynamics of persistence in school and school leaving, and the correlates of being at risk for not completing schooling;

5. Transition Patterns:
   a) the transition into high school
   b) the transition from high school into the world of work, postsecondary education, and family formation;

6. School Effectiveness: identification of those aspects of school structure and environment that are associated with positive student outcomes and school effectiveness; identification of those variables that best discriminate between effective and ineffective schools; and

7. Parental and Community Involvement/Social Capital.

The student questionnaire aims at collecting information that will furnish analysts and policymakers with data for evaluating student learning, persistence and other educational -- and eventually career -- outcomes. Since twelfth grade marks the culmination of the process of secondary schooling, most of the dynamic variables that were inquired into in the Base Year or First Follow-Up questionnaires were candidates for inclusion on the Second Follow-Up instrument. Priority was also given to items needed for cross-cohort comparisons with NLS-72 and HS&B. However, a further set of issues discussed during questionnaire development revolved around new items needed to capture transitional phenomena, in recognition that the Second and Third Follow-Ups constitute a new baseline for the further development of the study. Hence many of the new items that appear on the twelfth grade questionnaire are specifically designed to capture the transition of the non-college-bound to the labor force, and of others to colleges and other postsecondary institutions.

The field test student questionnaire comprises a locating section and six topical question series:

I. **School experiences and activities.** This section gathers information about the student's perceptions and attitudes about school, classes, teachers, homework, investment of effort to achieve good marks, school facilities, and disciplinary rules. This section will generate data for understanding, among other things the determinants of school completion and dropping out, and the effects of tracking/ability grouping on educational achievement and persistence.

II. **Plans for the future.** This question series repeats basic measures used in the Base Year and First Follow-Up. The data will allow analysts to assess how changes in plans and aspirations relate to educational experience, and how education and career plans are linked to future achievement.

III. **Language use.** This section provides data about student language use patterns and language ability variation.

IV. **Self concept and attitudes.** This section gathers information about student self-esteem, locus of control, achievement orientation, life goals, and attitudes.

V. **Work experience.** This section identifies the types and amounts of work that twelfth graders are engaged in after school and on weekends.

VI. **Family structure and environment.** The section on family structure and interactions inquires into household composition, intergenerational closure, family decision making, and the home education support system.

**Questionnaire Experiments in the Student Questionnaire:** An important objective served by the inclusion of methodological experiments in the field test is the identification and control of sources of nonsampling error. One major source of nonsampling error is nonresponse at either the respondent or the item level. Another source is reporting bias introduced through a number of factors associated with elements contained within the questionnaire, such as question form, question wording, response scales, position, or instructions, or with the mode of administration. These objectives figure prominently in the standards NCES has set for maintaining data quality in longitudinal surveys and we will focus on these through various methodological studies in the field test.

The NELS:88 Second Follow-Up field test provides opportunities for conducting a number of questionnaire experiments. The experiments described here were selected after a thorough review of the student instrument and a consideration of questionnaire items within the broader context of educational research was undertaken. In addition, the samples used in the different components of the study were considered, and experts in the area of questionnaire construction and survey methodological research were consulted. The experiments have two general purposes: (1) to help inform question construction for the main survey, and (2) to assess the impact on changes in responses of changes in questionnaire context from previous rounds.

Because of the large number of items used in the NELS:88 Second Follow-Up surveys, the number of item wording, item response format, and item context studies that are possible is staggering. For pragmatic reasons it was necessary to limit the number of such studies, but in such a way that the studies that are chosen are ones that will most benefit the NELS:88 program of research, in particular, and the field of educational research, in general, within the limits of available resources. With this in mind, we limited experiments to the student and parent questionnaires, with emphasis on the student questionnaire. This decision was driven by the primacy of the student in the NELS:88 research design, the large number of students and parents in the field test, and the complexity of the items on the student and parent questionnaires (often requiring sophisticated judgments and subjective reports). The decision to exclude the School Administrator Questionnaire from these studies was driven by the relatively small number of School Administrators in the field test sample, their consistently higher than average intelligence and/or education level, and the factual nature of most of the questions.

Three questionnaire development studies were conducted in the field test student survey, each addressing substantive and methodological issues. Studies one and two explore the effect of using different response formats for items in the Student Questionnaire. Study one examines the effect of using precoded categories on frequency estimates. Study two compares the "mark all that apply" response format with a format requiring respondents to give an explicit "no" response to a series of response categories. Study three looks at the effect of context on students' evaluations of their schools. Studies four and five concern the Parent Survey. See Section 6.1 for a discussion of the Parent component questionnaire experiments. See Section 3.4 for the results of the Student Questionnaire experiments.

### 2.2 Field Test Data Collection Procedures

The following section describes and evaluates the data collection procedures for the student component of the NELS:88 Second Follow-Up Field Test. The student component of the 2FU field test had as its primary purposes the testing of survey instruments, cognitive test items, and field procedures as those will be applied to 12th graders in the main study. Instrument development issues were discussed in the preceding section; cognitive test issues are elaborated in Section 3.10.2. This section is devoted entirely to student component data collection procedures; procedures for collecting data from dropouts, parents, and school administrators are addressed in sections 4.3, 6.3, and 8.2, respectively.

The field procedures implemented in the 2FU field test involved the same combined methods approach to data collection as did the NELS:88 1FU and HS&B. Specifically, two modes of survey administration were employed for collecting student data: 1.) an in-school, NORC-administered survey session, and 2.) an in-school, school coordinator-administered survey session. Due to the success of the First Follow-Up in collecting data in off-site sessions, it was not deemed necessary to again test these procedures for students attending schools which refused to participate; off-site sessions will again be employed in the 2FU main study for students attending schools which refuse to participate.

The combined methods approach to data collection has been assessed previously, as mentioned above. Our interest in testing these procedures yet again lay not in assessing the procedures for assessment's sake, but rather to detect any significant impact that a myriad of minor differences might have on the main study. For example, the 1FU was authorized to pay school coordinators an honorarium for their efforts in NELS:88 behalf; the 2FU is not so authorized. Gaining the participation of 10th graders in the second wave of a study may be easier or more difficult than obtaining second-semester seniors, with all the distractions of impending graduation, to participate in the third wave. These were the kinds of potential logistical problems that the 2FU field test was designed to assess.

In order to gauge the applicability of school coordinator-administered session procedures, seven schools that would ordinarily have been excluded from the field test effort (in that they refused both sample freshening and augmentation) were maintained in the field test sample of schools. The cluster of students in these schools ranged in number from one to nine. The remaining 87 schools had NORC-administered survey sessions.

### 2.2.1 In-School, NORC-Administered Survey Sessions

**Pre-Survey Day Procedures:** Once a school principal agreed to participate in the study, he or she was asked to select a date on which an NORC interviewer (also referred to as a team leader) could administer the Student Questionnaire and Cognitive Test to the schools' selected students. During this conversation, the school principal was also asked to appoint a school coordinator to act as a liaison between the school and NELS:88 2FU project staff. Most subsequent contact was between the NORC team leader and this school coordinator.

Approximately two weeks before the scheduled Survey Day, the team leader contacted the school coordinator to confirm receipt of NELS:88 materials at the school. The box contained: the School Coordinator Manual, student notification letters and permission forms, survey day invitations, Survey Day and Make-Up Day supplies, and the School Administrator Questionnaire. During this call, the team leader discussed the upcoming Survey Day activities including checking the enrollment of each student, monitoring the receipt of permission forms returned, and storage of the tests and questionnaires and asked the school coordinator to distribute the student notification/permission packets and the School Administrator Questionnaire immediately or at his or her earliest convenience.

The team leader again contacted the school coordinator approximately three days before the scheduled survey day session. The purpose of this call was to remind the school coordinator to distribute Survey Day invitations to the students, confirm the arrangements for the survey room, check on the enrollment and permission status of each student invited to participate and coordinate Survey Day plans. The verification of enrollment status primarily pertained to the determination of whether any selected students had dropped out or transferred out of the school, and hence become ineligible for data collection. Transfer students were not pursued further.

If a student's parent had denied permission for participation in the study, the team leader attempted a conversion of the refusal in a telephone call to the household. Telephone rather than letter conversions were employed because there was insufficient time before Survey Day for sending such a letter by regular mail.

**Survey Day Procedures:** On Survey Day, the team leader and clerical assistant arrived at least 30 minutes prior to the scheduled start of the session. The team contacted the school coordinator and began preparation of the room and materials for the session. Ideally, the survey room was a space separate from other activities and contained more seats than were required for the total number of students, a desk or table for the team leader and clerical assistant, a blackboard, and a clock. Classrooms, libraries, and cafeterias were selected by the school staff for the session; this was primarily due to the large number of students (35 - 40) we asked that the room accommodate. Libraries and cafeterias met most of the criteria but were less than satisfactory for our purposes because of the distractions present in these multi-use spaces.

The session generally began with a roll call (although some teams preferred to take names as the students entered the room). The students were seated randomly with a desk or chair between them (when possible). When it appeared that most students had arrived, the team leader gave the list of those missing to the school coordinator. The school coordinator introduced the members of the team and then left to locate the missing students. The team leader began the session by reading a script that included a welcome, an introduction to NELS:88, statements about confidentiality and the voluntary nature of the study, and about the Student Questionnaire. Together with the students, the team leader then reviewed the instructions for filling out the questionnaire booklet.

(The paragraph above describes how the ideal-type session began. Not all school coordinators were present for the beginning of the session, nor did all search for missing students. When the school staff could not provide this level of support, the team leader requested that the public address system be used to remind missing students of the session and its location. If this was not possible, the team leader sought out missing students (if information was available to do so) after starting the students present on the questionnaire. In this circumstance, the clerical assistant served as overseer of the group but was instructed not to answer questions regarding the questionnaire.)

Students were given sixty minutes after the conclusion of the script to complete the Student Questionnaire. (Students included in data collection through sample freshening were also instructed to complete a New Student Supplement during this time peroid.) Team leaders were instructed not to allow additional time on the questionnaire, even if some students had not finished. This rule was instituted so that questionnaire length could be accurately assessed in the field test and refined for the main study. The team leader answered students' questions at their desks. All questionnaires were collected at the end of the sixty-minute period.

In order to replicate the survey experience across schools, Survey Day procedures were standardized. For example, all team leaders read the same script explaining the purpose of and procedures for the study. In addition, all team leaders were instructed to answer students' questions about the Student Questionnaire in the same manner. If a student inquired as to what a particular item meant, team leaders were instructed to read the question aloud to the student rather than offer their own interpretation of the item. If this did not satisfy the student, he/she was told to answer the question based on what they thought the question was asking. Team leaders completed a Tally Sheet indicating which items engendered inquiries from students and reported in detail on especially problematic questions. (The number of inquiries for each item is indicated in the Student Questionnaire in Appendix E.)

At the end of the Student Questionnaire administration, a transcript permission form was distributed and students were asked to complete the form and place it inside the back cover before handing in the questionnaire. (Less than one percent -- 0.62 -- of students refused transcript release in the field test.) Students were then given a ten-minute break before the start of the cognitive test battery. Usually some type of refreshment, such as a juice box and granola or candy bar, was served during this break. After students had an opportunity to snack and stretch, the team leader reconvened the session, read the script introducing the cognitive test battery, distributed the test booklets, and began the first test section.

Students were randomly assigned one of five forms of the timed Cognitive Tests. In order to assess the validity and reliability of a pool of multiple choice and free response items for four subject areas -- mathematics, science, reading, and history/civics/geography -- within an 85-minute testing period, field test participants were tested in two of the four subject areas. Each test booklet began with a short series of questions eliciting basic demographic information; students were given five minutes to complete these questions. There were three major sections of the test period; the first was 25 minutes and the second was 17 minutes. The third section contained four free response items; students were given 12 minutes to complete each of these and the subsequent short questions assessing the adequacy of time, the difficulty of the question, and the level of effort they expended on the item. The cognitive test battery and results obtained are discussed in detail in Section 3.10.

One complication in test administration arose during data collection. NORC field test staff received feedback from the field managers that the third (free response) section of the cognitive test battery was posing problems in the schools. In many schools, the 12-minute time period for each question was too long and students became restless and disruptive. Before this came to the attention of project staff in Chicago, team leaders improvised by allowing students to continue work on any uncompleted sections of their Student Questionnaire (and performing the edit and retrieval after each student finished). This practice was stopped as soon as it came to our attention. Team leaders also occasionally allowed students to read books and/or do homework in an effort to maintain an orderly and quiet testing environment. Our recommendations to control this type of improvisation involve more thorough training of NORC field representatives on the general principles of cognitive test administration and classroom control techniques.

While students were completing the Cognitive Test, the clerical assistant and team leader edited the Student Questionnaires, checking that critical items were completed in full. If data were missing, the team leader attempted retrieval at the student's desk when he/she had completed a test section. At the end of the testing session, students were instructed to close and hand in their test booklets. The school coordinator or team leader dismissed the students to their classes after expressing NORC's appreciation for their participation.

**Post-Session Procedures:** Following the dismissal of the students, team leaders assessed the need for a Make-Up Day session with the school coordinator. Because the field test budget provided for only 12 NORC-administered Make-Up Days and the turnout for Survey Day sessions was often at about fifty percent, this meant a delicate negotiation between the team leader and the school coordinator who would be responsible for conducting the Make-Up Day session. Many schools did not cooperate, citing overburdened staff, understaffing, and the level of effort requested of the school coordinator. In those schools that agreed to conduct Make-Up Day, as well as in the sessions conducted by team Leaders, the same procedures implemented on Survey Day were also employed for the Make-Up Day session. Before leaving the room, the team leader and clerical assistant packaged completed materials for shipment to Chicago and cleaned up the survey room. The team leader also inquired about the status of the School Administrator Questionnaire, collected it if complete, and personally thanked the school principal and the school coordinator for their cooperation and assistance.

### 2.2.2 In-School, School Coordinator Administered Sessions

Seven schools refused both sample freshening and augmentation and thereby provided us with the opportunity to field test the procedures for school coordinator administered Survey Day sessions. Of these seven schools, one school refused during the data collection period to participate.

Although school coordinators were responsible for administering the Survey Day session, team leaders still performed other pre-session tasks like confirming receipt of data collection materials and monitoring the status of enrollment and permission. To assure comparability between the two modes of data collection, a School Coordinator Manual was prepared detailing the procedures to be used. This manual was distributed to all school coordinators (regardless of who was responsible for conducting the session), but the team leader reviewed this manual more closely with school coordinators collecting data.

In essence, the School Coordinator Manual described procedures identical to those utilized by team leaders, with several notable exceptions. In order to maintain confidentiality and reduce response bias (Sudman and Bradburn, 1982), school coordinators were instructed not to do anything that would require them to look at the students' questionnaires. This procedural rule was specified as crucial to the success of the study, and was repeated throughout the manual. School coordinators were warned that the integrity of the data might be compromised if students believed that school staff were privy to their responses.

More specifically, school coordinators were instructed not to respond to students' questions about the questionnaire or cognitive test. In the introductory script read by the school coordinator, students were told that questionnaire items could not be clarified and that they should answer according to what they thought the question meant. Similarly, school coordinators could not perform the edit and retrieval of critical items due to this restriction. Rather, students were instructed to go back through their own questionnaire and make sure that they had answered every question. Students were instructed to place their questionnaires and tests in a confidential envelope and seal it before turning it in to the school coordinator.

### 2.2.3 Quality Control and Retrieval of Critical Items

The Student Questionnaire contained 21 items designated as critical to the research concerns of the study. As such, an on-site edit and retrieval of these items, if missing, was considered necessary. The critical items in the Student Questionnaire and the substantive content of each item is presented in Table 2.2.1. This section presents a brief evaluation of the results of the editing and retrieval, specifically with regard to identifying problematic questions. A high retrieval rate for an item often indicates a wording or formatting problem.

As mentioned above, while students were completing the Cognitive Test, the clerical assistant and team leader edited the Student Questionnaires, checking that critical items were completed. If data were missing, the team leader attempted retrieval at the student's desk when he or she had completed a test section. Interviewers were instructed to perform the edit only on completed portions of the questionnaire. In other words, interviewers did not retrieve on a critical item if the respondent had stopped answering questions before that point in the questionnaire. The percentage of students (of the 996 whose questionnaires were data entered) from whom data retrieval was attempted is displayed next to the question number in Table 2.2.1, as well as in Appendix E.

There was no attempt to evaluate the effectiveness of the edit and retrieval procedures. Experience in both the Base year and First Follow-Up suggest that this assessment is unnecessary. The results of the Base Year analysis indicated that the in-school edit was overwhelmingly successful in assuring high quality data in cases where students possessed relevant information (Ingels et al., 1987).

## 2.3 Data Preparation Procedures

Data preparation encompasses all activities associated with converting data from completed Student Questionnaires to data files. These included receipt, manual edit, coding and data entry.

**Table 2.2.1 Retrieval Rate for Critical Items in the Student Questionnaire**

| Q. NO. | Retrieval Rate (%) | Text |
|---|---|---|
| 1. | 1 | Please print your name, address, and telephone number. |
| 1A. | 16 | What is your Social Security Number? |
| 2. | 3 | Is your mother's address and telephone number the same as yours? |
| 3. | 5 | Is your father's address and telephone number the same as yours? |
| 4. | 21 | Please write the name, address, and telephone number of a relative or close friend who does not live with you. |
| 5. | 1 | What grade are you in? |
| 9. | 3 | When was your last unexcused absence from high school, if any? |
| 11. | 3 | Which of the following best describes your present high school program? |
| 24. | 8 | Overall and in the following subjects, about how much time do you spend on homework EACH WEEK, both in and out of school? |
| 40. | 2 | How important is each of the following to you in your life? |
| 41. | 2 | How sure are you that you will graduate from high school? |
| 45. | 4 | As things stand now, how far in school do you think you will get? |
| 46. | 8 | Have you taken or are you planning to take any of the following tests this year? |
| 56.A | 6 | If you go on to school after you leave high school, will you most likely... |
| 61. | 14 | How many schools have you applied to? |
| 66. | 19 | Which of the categories below comes closest to describing the job or occupation you expect or plan to have right after you graduate from high school and when you are 30 years old? |
| 93. | 8 | Do you have any children of your own? |
| 102. | 11 | When did you last work for pay, not counting work around the house? |
| 109. | 9 | Which of the following people live in the same household with you? |
| 112. | 17 | In a typical month, how many school days do you miss because of taking care of your own child or your brothers and sisters? |
| 113. | 14 | Lots of things happen in families that affect young people. In the last two years, have any of the following things happened to your family? |

Source: NELS:88 Second Follow-Up Field Test Data

**Receipt Control:** Team leaders were instructed to return Student Questionnaires and Cognitive Tests to NORC after all Survey Day activities at the school had been completed. Federal Express was used, to ensure the arrival of the questionnaire and test package and to expedite processing of the materials.

Questionnaires and tests were receipted in the Receipt Control Shop of the NORC Lake Park Facility. Upon delivery of the package, a receipt control clerk reconciled the enclosed questionnaires and tests with the transmittal form completed by the team leader to identify any discrepancies between what was thought to have been sent and what was actually received.

Receipt control clerks updated dispositions in the Survey Management System. (The SMS is discussed below). Separate SMS variables were used to track questionnaire and test status. The current questionnaire and test dispositions were entered directly from the completed instruments, in a batch update mode by type of document and disposition. As a final step, the main disposition for each student case was updated from the transmittal form completed by the team leader. Questionnaire and test dispositions were again updated when questionnaires were sent to data entry and tests were shipped to the Educational Testing Service (ETS).

Receipt control clerks also updated the school SMS record, entering the current main school disposition from the Survey Day transmittal form and the school questionnaire disposition from the completed instrument. Administrative variables such as survey date and make-up date were also updated.

**Survey Management System:** One of the goals of the 2FU field test was to rigorously test the feasibility of using a relational database for the 2FU Main Study Survey Management System (SMS). Relational databases are sets of related information stored in tables. Tables within a database can represent widely varying entities. However, no matter how different the information contained in tables is, information from those tables can be linked, as long as a logical connection between the tables can be established. This "linking" feature of relational databases is particularly well-suited to managing data from multiple sources where the sources are linked in a hierarchical fashion. NELS:88 survey management data is structured in this way.

During each round of NELS:88, survey management data are collected for each of the various study components -- student, school, district, teacher, and parent -- and entered into the SMS. In the past, data for each component were kept on separate Survey Management Systems, making any linking of data from different components impossible within the system. In previous NELS:88 field tests, SMS data bases were constructed and maintained by project staff. Because of the availability of mature relational data base software and the plan to implement an SMS in such a program for the 2FU main study, the decision was made to develop and test the system during the field test.

Second Follow-Up field test SMS data from the various components were housed in separate tables within a single relational database, thus allowing the linkage of data from different components within the system. Overall, the SMS developed for the field test was judged to be a powerful management tool that was generally easy to use. The 2FU field test SMS database was created and manipulated using Structured Query Language (SQL). The syntax of SQL is quite simple, and because of this, supervisors and other users were able to query the system with relative ease. Much of the SMS program for the main study will be adapted from the field test system.

**Editing/Coding:** Editing was performed in two steps. During the administration of the Cognitive Test on Survey Day, the clerical assistant performed a critical item edit of the Student Questionnaires and the team leader attempted to retrieve missing data. See Section 2.2.3 for further details of this scan edit.

In-house editing was performed at NORC's Lake Park Facility. A total of 996 questionnaires (approximately 250 of each form) were edited for missing items, using a coding system consistent with that applied to the NELS:88 First Follow-Up Field Test data. Marginal indications of "Refused" or "Don't know" (whether by the team leader or the student) were coded into standard reserve codes (as "7" and "8", respectively). Additionally, unmarked missing items, whether skipped legitimately or illegitimately, were marked "9" for missing, and a "6" was used to indicate an unacceptable multiple response. Open-ended questions were coded "1", to indicate that data were present. For some items, these reserve codes were preceded by a "9", depending on the number of columns corresponding to the question. For example, a missing item with response codes ranging from 00 to 12, with a marginal indication of refusal, was coded a "97".

**Data Entry:** Two file layouts (essentially column position and field length for every item in the questionnaire) were developed for data entry of the questionnaires, one for forms A and B, the other for forms C and D. NORC's data entry subcontractor, BSI, prepared a test file of the data for review, keying five of each of the four forms.

The remainder of the 996 edited and coded Student Questionnaires were then sent for data entry. Questionnaire data were converted to machine-readable form using a conventional key-to-disk method. Data were fully verified (that is, re-entered independently by a different operator) with the software comparing all fields for an exact match. No machine cleaning of the data was performed, to facilitate the identification and analysis of problems within the questionnaire.

## 2.4 Recommendations for Changes in Student Data Collection Procedures

Notwithstanding the critical importance of the sample design and the development of effective survey instruments, securing the full participation of sample members and collecting data from them is the major activity of the NELS:88 Second Follow-Up. For a given set of survey instruments, the data collection plan is the key to controlling nonsampling error. To accomplish the study's research objectives, response rates for all samples must be as high as feasible. The sophistication and efficiency of a chosen sample design can be vitiated by low completion rates for the whole or any population subgroups. And needless to remark, a longitudinal study hinges upon the collection of accurate follow-up data and locating information for future surveys.

The Second Follow-Up Field Test was designed as a trial of the basic data collection plan for the main study, within the special constraints occasioned by the need to modify the field test design to accommodate special cognitive testing needs. Given that in general securing the participation of twelfth graders is thought to be more difficult than gaining the cooperation of eighth or tenth graders, the major data collection challenge of the 2FU student survey will be to achieve a high student response rate. This will almost certainly entail eliciting the fullest cooperation from the school coordinator, as well as directly motivating members of the student sample. In the course of testing the design and operational procedures of the main study, many problems--small and large--were identified in the field test. This section addresses several of the major lessons we consider important to carry forward to the preparations for the 1,500 in-school data collection sessions scheduled for February through May, 1992.

**School Coordinator Role:** The role of the school coordinator is a very important one. The success of NELS:88 depends in large measure on the levels of cooperation and enthusiasm that such individuals bring to their responsibilities. If the school coordinator believes in the value of the research project and devotes the necessary time and attention to the task, the Survey Day is more likely to be a success. For example, a school coordinator can maximize attendance at Survey Day (thereby increasing achieved response rate and minimizing the level of effort necessary and associated costs) by "rounding up" students who do not report for the session at the scheduled time. Increasingly, however, all school staff members are overburdened with day-to-day responsibilities and the problems resulting from overcrowding and austere budgets; in addition, schools are asked more and more often to administer tests and to host research projects.

NELS:88 demands much of the staff person appointed to this position by the school's principal. Following are several recommendations to ameliorate the burden imposed on NELS:88 school coordinators.

a.) **Seek approval from OMB to reinstate the honorarium.** In previous waves of NELS:88, school coordinators were offerred a modest honorarium to partially compensate them for their time and efforts and to serve as a tangible mark of appreciation. Because in the NELS:88 1FU we were unable to meet the OMB requirement that a statistically significant effect of honoraria on school's propensity to agree to the study be demonstrated, no honorarium was proposed for the 2FU. This situation is particularly difficult because of the overlap between 1FU and 2FU schools combined with the proclivity of school principals to appoint the same person to the school coordinator position, and the fact that the 2FU includes two components new to NELS:88 (the course offerings and enrollment and academic transcript components) which are especially burdensome for school staff. (See Section 10 for further details about these components.)

b.) **Decrease the scope and complexity of school coordinator responsibilities.** There are several simple changes in procedures that can decrease the burden of NELS:88 on school staff. For example, we recommend that questionnaire distribution to selected teachers and the school principal be handled through direct mail to those individuals, focusing the school coordinator's responsibilities on the pre-survey day, in-school, and school records data collection activities. Similarly, we recommend that mailings to the school coordinator of informational and survey day materials be increased in number but decreased in size and complexity. These simple, low-cost measures are anticipated to reduce the perception that fulfilling the liaison role to NELS:88 is an overwhelming and time-consuming task.

c.) **Attempt to engage the school coordinators in the project from our first contact with them.** In the past, we have relied solely on our experienced and persuasive field staff to convince school coordinators of the importance of NELS:88, to motivate them to work hard in our behalf, and to communicate details of their responsibilities at appropriate points in time. For the 2FU Main Study, we recommend that a newsletter or brochure be prepared this summer for distribution to the school coordinators during the interviewer's fall visit to the schools to freshen the sample, identify the selected teachers, etc. This document would contain some (carefully phrased) news of 1FU results, reiterate the significance of NELS:88 within the body of education research, and generally "sell" the school coordinator on committing his or her time and energy to the project. Additionally, interviewers would be strongly encouraged to schedule a brief meeting with the appointed school coordinator during their fall school visit so their 2FU relationship begins on a strong foundation.

d.) **Provide better information to facilitate performance of school coordinator responsibilities.** In addition to an attractive and motivational newsletter, we recommend providing school coordinators with more information about their specific responsibilities and the timetable in which those duties will be performed during the interviewer's fall visit to the school. This should include a description of our planned contacts with them, as well as of the tasks that they will be asked to accomplish, to enable them to better plan for these assignments.

One particular group of school coordinators -- those in schools with small clusters of students -- will be asked to take on the responsibility of conducting all data collection activities. These school-administered sessions have always been a necessity to contain costs; yet, ensuring the consistency of procedures across survey sessions is much more difficult when school staff are the data collectors, and nearly impossible when the school coordinator has not been adequately motivated nor provided with user-friendly instructions.

For these school coordinators, better materials are essential. Apart from motivating them to conduct the Survey Day and Make-Up Day sessions through an intellectually engaging newsletter and symbolic compensation, we must equip them with the materials to fulfill these responsibilities well. We recommend that administration materials (specifically, the School Coordinator Manual) be made more clear, concise, and engaging.

**Renewing the Commitment of Longitudinal Cohort Members:** Our experiences in the 2FU Field Test have served to remind us all of the importance of the commitment instilled in our sample members. At this time in their lives -- on the eve of high school graduation and entry into a world of greater freedoms and responsibilities -- we must work especially hard to renew their dedication to NELS:88.

To demonstrate that we are not taking them for granted, we recommend that the timing and nature of contacts with sample members be altered. We suggest a postcard be distributed to sample members via the school coordinator immediately subsequent to the interviewer's fall visit to the school. The purpose of this postcard is simply to remind each of them that we will soon be contacting them with arrangements for data collection.

For 1FU and 2FU freshened students, a letter emphasizing the nature and importance of the study would accompany the postcard. We recommend this special attention to freshened students because of our experiences in the 1FU Main Study and 2FU Field Test, which give clear evidence of the disinclination of new cohort members to participate.

In addition to the postcard, the distribution of a student newsletter is suggested for two weeks before the scheduled survey day session. This newsletter, designed to be attractive and engaging to this age cohort, will contain news of the study (carefully selected and phrased results from the 1FU, for example) and a reminder of the importance of their participation. This newsletter will be followed by the student notification letter, letter to parents, and permission form about one week before data collection and the survey day invitation one to two days before the scheduled survey day session.

## Chapter 3: Analysis of Student Questionnaire and Test Data

This chapter discusses the results of in-depth analyses of the Student Questionnaire and Cognitive Test data. Student response rates -- overall, by school sector, by state, and by participation in previous waves -- are briefly examined. In subsequent sections, specific recommendations are offered for improving the effectiveness of the questionnaire in collecting high quality data, based on (a) item nonresponse analysis, (b) logical consistency analysis of filter and dependent questions, and (c) analyses requested by the Office of Management and Budget in correspondence relating to the clearance document.

The results of questionnaire experiments -- investigating effects of response format (open - v. closed-ended), response order, and item order -- are also discussed and analyzed, and revisions to specific items based on these analyses are proposed. The reliability of two attitude scales used in the Second Follow-Up Field Test Student Questionnaire, which were also used in the First Follow-Up and Base Year versions of the instrument, are estimated. In the final questionnaire section, general recommendations for changes in the instrument are presented. No assessment of the New Student Supplement was conducted, due to the very small number of completions (six) and due to the fact that the instrument is virtually unchanged from its 1FU form.

The final chapter section focuses on the analysis of the Cognitive Test Battery data. The objectives and structure of the cognitive tests used in the Base Year and First Follow-Up studies are discussed, followed by an analysis of the data collected in the multiple choice and free response sections of the tests. In conclusion, recommendations for the main study tests are submitted.

### 3.1 Student Response Rates

In this section, rates of student participation in the NELS:88 Second Follow-Up Field Test are analyzed. Response rates were computed by dividing the number of students who participated in the field test by the number of students with whom data collection was attempted. "Participation" was defined as completion of a cognitive test battery, the primary focus of the 2FU field test. Types of sample members excluded from the denominator include: those who attended a school (or district) that refused to participate (141), early graduates (48), those who transferred to another school after the sample was finalized in December (42), those attending the two schools which refused during data collection (40), those with a permanent incapacity (14), those who moved out of the field test area or the country (28), those classified as unlocatable during tracing (51), independent and at-home-study students (9), and those deemed ineligible because of a language barrier (1). Of the total 3,704 sample members, 374 were declared to be ineligible for data collection as just described, 3,222 were eligible for student data collection, and 108 were eligible for dropout data collection. Figures 3.1.1, 3.1.2 and 3.1.3 display student response rates by sample type (core students versus freshened and augmented students), field test state, and school sector.

Overall, 70 percent of all students with whom data collection was attempted participated in the study. A total of 2,252 completed Student Questionnaires and 2,254 cognitive test were received at NORC. (Note: A number (184) of cognitive tests were received too late for ETS to utilize the data.) The percentage of Catholic students participating in the field test was substantially higher (85 percent) than the percentages of public school and private/other students who participated in the study.

There were considerable differences in rates of participation for the five field test states. At opposite ends of the spectrum were Texas which had an average completion rate more than ten percentage points higher than the overall average, and California, which had an average completion rate nearly ten percentage points below the overall average.

Participation in the field test appears to be strongly associated with respondent sample type. While response rates for core students and students who were augmented for the 2FU Field Test were roughly the same as the overall average, response rates for students added to the sample through freshening in the first and second follow-up studies were much lower than the overall average. The participation rate for 2FU freshened students is of particular note - it is more than 20 percentage points lower than the overall average. It will be important in the main study to communicate with these freshened students and to convince them of their importance to the study in an effort to elicit higher levels of participation from this population.

Far higher response rates could have been obtained, though at a greater cost to budget and schedule and without necessarily maximizing the purposes of the field test, which are rather different from those of the main study. As noted above, the primary objective in the student survey was to amass a large number of cognitive test observations at an early point in the schedule, rather than to go back to schools for repeated make-up days in order to achieve a higher response rate over a longer schedule. The fact that over two-thirds of the 2FU field test sample comprised students new to NELS:88 -- and the consideration that their effect on the participatory propensities of students who were part of the longitudinal cohort is unknown -- limits severely the generalizibility of any findings on tendencies to respond. Nevertheless, we are well aware that second-semester seniors are, as a group, notoriously more difficult to survey than are younger students and ways to renew the motivation and commitment of longitudinal sample members in 1992 must be identified and incorporated into main study procedures.

### 3.2 Item Nonresponse Analysis

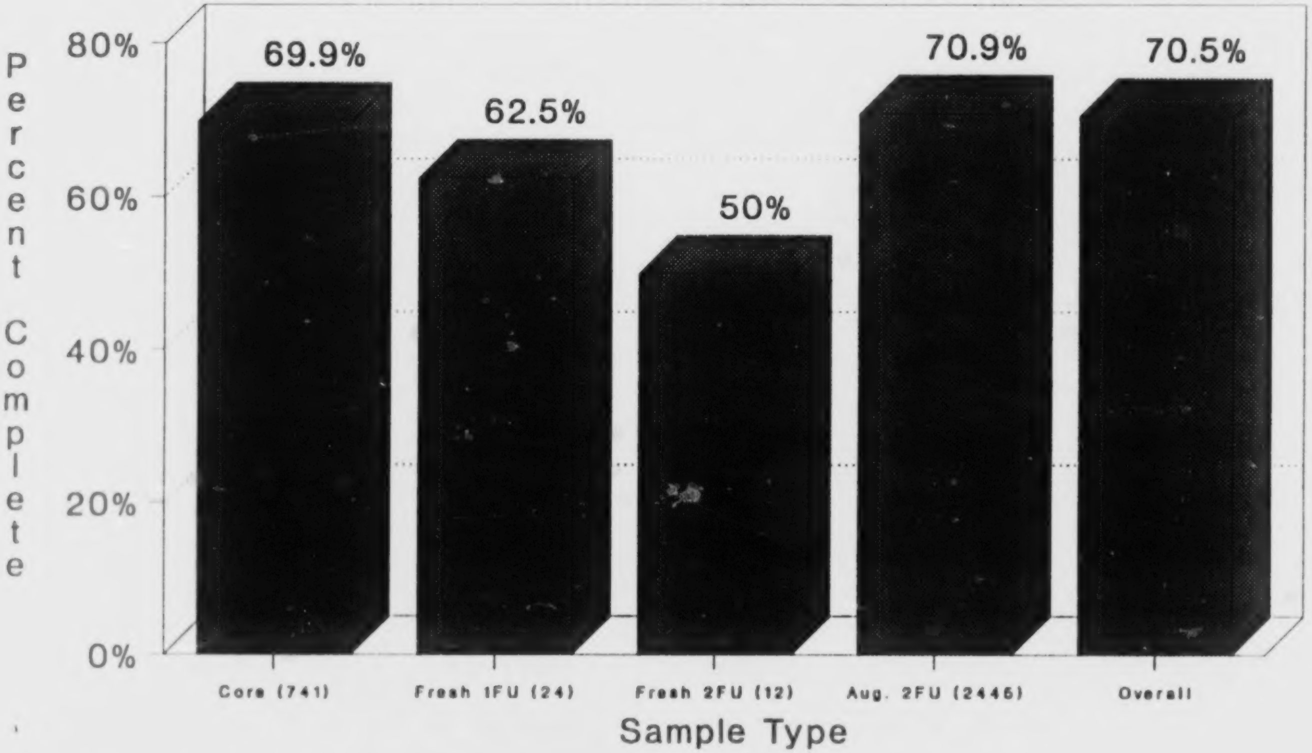
In this section we present an analysis of item nonresponse in the Student Questionnaire. Item nonresponse may indicate that a question is poorly formatted, not being understood by the respondent or asks for information the respondent either cannot recall at the time he/she is completing the survey or that he/she simply does not possess. In addition to identifying items with high nonresponse rates, we present recommended wording and format changes likely to improve response rates.

Because of the complexity and length of the questionnaire, meaningful item nonresponse rates are no simple matter to calculate. In our analysis we attempt to account for missing responses to filter questions when assessing item nonresponse to dependent questions. In addition, we attempt to separate nonresponse in later sections into that which is a function of the item and that which is a function of lack of time to complete the questionnaire. Where applicable, we explain our techniques and give justifications for them.

**Overview of the results.** Before we begin our item-by-item analysis, it is useful to state the lessons we have learned. We have found five major sources of item nonresponse in the Student Questionnaire: (1) nonresponse due to poorly designed questions, (2) nonresponse due to overcrowded questions, (3) nonresponse due to difficult, vague, or abstract language, (4) nonresponse due to difficult judgments, and (5) nonresponse due to the inapplicability of the item to respondents. An additional type of nonresponse is that which is explainable by responses to prior items (e.g. legitimate skips). Items falling into this latter category are mentioned in our report for the sake of completeness, even though they do not constitute a problem.

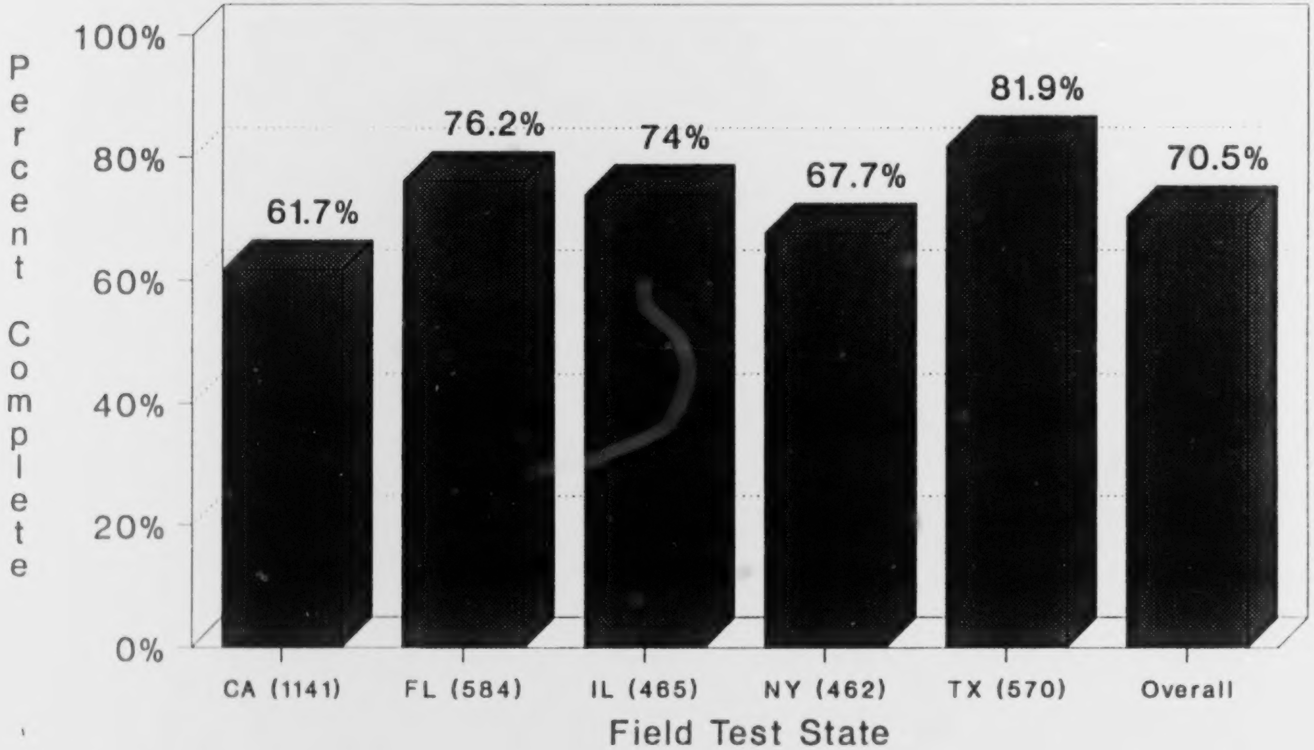
One particularly salient example of the first source of item nonresponse is the practice of embedding a skip item as the first response in a question matrix. Examples of this are found in Q16, Q19, and Q22. In general, we recommend against this practice for a number of reasons discussed below, and suggest instead to treat the skip question as a separate item. Another example is placing a filter item just before a page turn.

Figure 3.1.1. Student Completion Rates by Respondent Sample Type



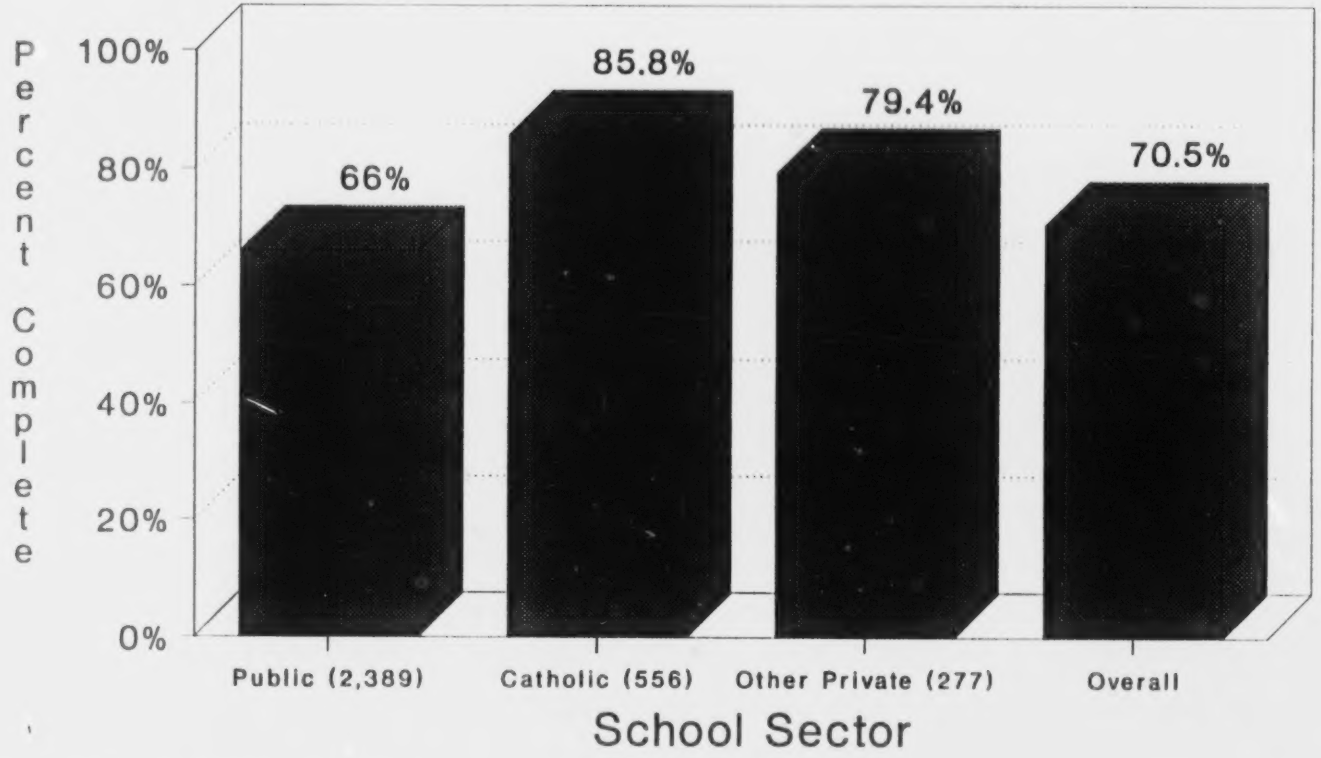
Base: 3,222 students eligible for
data collection
Source: NELS:88 2FU Field Test Data

# Figure 3.1.2. Student Completion Rates by Field Test State



Base: 3,222 students eligible for
data collection
Source: NELS:88 2FU Field Test Data

## Figure 3.1.3. Student Completion Rates by School Sector



**Base: 3,222 students eligible for data collection**
**Source: NELS:88 2FU Field Test Data**

An example of the second source of nonresponse can be found in the question about postsecondary institutions to which the respondent has applied (Q61). The question format for Q61 is so crowded that it is difficult for respondents to tell which parts of it they should fill out. For this and other high nonresponse items, we suggest a simplified format. Question 22 provides a good example of the third source of nonresponse. Definitions of key terms and wording simplifications are suggested as ways of reducing nonresponse.

Asking respondents to make judgments such as "on the average" (Q111) and "in a typical month" (Q112) is likely to elicit the fourth type of nonresponse because of the difficulty people have in making these kinds of estimates. In addition, difficult judgments such as these may not only lead to nonresponse, but may lead to poor responses (e.g., wild guesses) among respondents rendering the judgments. Finally, the part of Question 94 asking for the month and year of the birth of the respondent's second child is an example of the last type of nonresponse. Only a very few respondents had one child, and none had two.

Item nonresponse rates were calculated in three steps. First, dichotomous variables were created for each questionnaire item with "1" indicating no response was given for the item and "0" indicating a valid response was given for the item. Second, in order to make all items comparable, the dichotomous variables were adjusted so that item nonresponse rates were calculated only for those respondents for when a response was legitimate or applicable. Finally, item nonresponse rates were derived by calculating the mean of the dichotomous variables. The resulting item nonresponse rates indicate the proportion of respondents eligible to answer an item who failed to answer the item. The item nonresponse rates were calculated for the 996 students whose questionnaires were data entered.

Upon initial review of the item nonresponse rates, it was discovered that a considerable number of respondents did not reach the last four sections of the questionnaire. This had the effect of increasing nonresponse rates for items in those sections (A discussion of nonresponse rates by item position can be found in section 3.5). The goal of this portion of the analysis is to identify item nonresponse problems that are related to factors such as question format, question content, and skip instructions. Because of this, an attempt was made to "control" the impact of questionnaire length on item nonresponse rates so that the effects of the factors mentioned above could be more clearly ascertained. In order to control for the effect of questionnaire length, adjusted nonresponse rates were calculated for items within the last four questionnaire sections. The adjusted rates were calculated using the following steps. First, respondents' answers to all questions in the last four sections of the questionnaires were checked. Second, respondents who missed all items within a section were removed from the base of respondents used to calculate nonresponse rates for items within that section.

In selecting items for analysis, items were first divided by questionnaire section. Within each questionnaire section, items were sorted by item nonresponse rate in descending order, and all items with an item nonresponse rate of .20 or greater were chosen for analysis. Item nonresponse rates for items meeting these criteria in the Student Questionnaire are displayed in Table 3.2.1.

### Questionnaire Section 1: Address Information

Section 1 consisted of locator information. Items were not keyed; consequently, they were not analyzed for nonresponse.

## TABLE 3.2.1

### Student Questionnaire Item Nonresponse [1]

| ITEM | Multiple Response | Refusal | Don't Know | Missing | TOTAL [2] | ADJUSTED [3] |
|---|---|---|---|---|---|---|
| Q14B | N/A | N/A | N/A | 0.34 | 0.34 | 0.34 |
| Q16 filter | 0.00 | 0.0 | 0.0 | 0.92 | 0.92 | 0.92 |
| Q19 filter | 0.00 | 0.00 | 0.00 | 0.95 | 0.95 | 0.95 |
| Q22 filter | 0.00 | 0.00 | 0.00 | 0.37 | 0.37 | 0.37 |
| Q22A | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 | 0.24 |
| Q22B | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 | 0.24 |
| Q22C | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 | 0.24 |
| Q22D | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 | 0.24 |
| Q22E | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 | 0.24 |
| Q22F | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 | 0.24 |
| Q49 | N/A | N/A | N/A | 0.33 | 0.33 | 0.33 |
| Q50A | 0.01 | 0.00 | 0.00 | 0.24 | 0.25 | 0.25 |
| Q50B | 0.01 | 0.00 | 0.00 | 0.27 | 0.28 | 0.28 |
| Q51 | N/A | N/A | N/A | 0.25 | 0.25 | 0.25 |
| Q52 | 0.00 | 0.00 | 0.01 | 0.40 | 0.40 | 0.40 |
| Q53 | 0.00 | 0.00 | 0.00 | 0.31 | 0.31 | 0.31 |
| Q54 | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 | 0.24 |
| Q55 | N/A | N/A | N/A | 0.50 | 0.50 | 0.51 |
| Q61C | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 | 0.22 |
| Q61D | 0.00 | 0.00 | 0.01 | 0.37 | 0.38 | 0.37 |
| Q61E | 0.00 | 0.00 | 0.00 | 0.42 | 0.42 | 0.42 |
| Q61F | 0.00 | 0.00 | 0.00 | 0.45 | 0.46 | 0.46 |
| Q66A | 0.00 | 0.00 | 0.09 | 0.20 | 0.29 | 0.29 |
| Q66B | 0.00 | 0.01 | 0.06 | 0.14 | 0.20 | 0.20 |
| Q75 | N/A | N/A | N/A | 0.89 | 0.89 | 0.38 |
| Q76 | N/A | N/A | N/A | 0.58 | 0.58 | 0.34 |
| Q94 (1ST) | N/A | N/A | N/A | 0.96 | 0.96 | 0.92 |
| Q94 (2ND) | N/A | N/A | N/A | 1.00 | 1.00 | 1.00 |
| Q95 | 0.00 | 0.00 | 0.00 | 0.91 | 0.91 | 0.83 |
| Q96 | N/A | N/A | N/A | 0.97 | 0.97 | 0.94 |
| Q111 | 0.00 | 0.00 | 0.00 | 0.57 | 0.57 | 0.24 |
| Q112A | 0.00 | 0.00 | 0.00 | 0.46 | 0.46 | 0.07 |
| DATE | 0.00 | 0.00 | 0.00 | 0.46 | 0.46 | 0.46 |

Source: NELS:88 Second Follow-Up Field Test Data.

---

[1] The nonresponse categories are defined as follows: Multiple response -- respondent circled more than one response to an item in which only one response was allowed; Refusal -- respondent indicated he or she refused to answer the question either by writing directly on the questionnaire or by telling the interviewer; Don't know -- respondent indicated "don't know" either by circling a "don't know" response option, by writing "don't know" directly on the questionnaire, or by giving a "don't know" response to the interviewer; Missing -- respondent failed to give any response to a question that was not part of a skip pattern.

[2] Total nonresponse is the sum across all nonresponse categories.

[3] Adjusted nonresponse deviates from total nonresponse only for the last four sections of the questionnaire. Adjusted nonresponse was calculated by excluding respondents who had skipped an entire section from both the denominator and the numerator of the item nonresponse ratio.

Questionnaire Section 2: Your School Experiences and Activities

**Analysis: Q16, Q19, and Q22**

Nonresponse rates above twenty percent were found for ten items. Three of the items were filters embedded in a complex matrix question. These three are (1) "Have not taken a science class in the last two years" (embedded in Q16 "In your current or the most recent science class you took within the last two years, how often do/did you...?"), (2) "Have not taken a mathematics class in the last two years" (embedded in Q19 In your current or most recent math class you took in the last two years, how often do you...?"), (3) "Have not taken a vocational course" (embedded in Q22 "In your current or most recent vocational course you took within the last two years, how much emphasis did/does you teacher place on the following objectives?")

Because of the ambiguity in the response format, it is impossible to tell whether a nonresponse means a negative response to the item (i.e., the respondent had taken one of the types of courses indicated in the question) or a missing response, without considering response to subpart A through L. In general, we feel that this kind of ambiguity should be avoided. The problem is complicated by the fact that this filter comes in between an introduction and related questions on a different topic. Thus, the natural flow and logical connection between two important parts of the question is interrupted.

**Recommendations:** For each of these items we recommend that the filters be placed as separate questions before Q16, Q19, and Q22, respectively. Also, we recommend that both the affirmative and the negative response categories be included in the filter, each with appropriate branching instructions.

Proposed revisions to these filter questions are listed below.

16. Have you taken a science class in the last two years?

Yes . . . . . . . . 1 **(GO TO Q.16A)**
No . . . . . . . . 2 **(SKIP TO Q.19 ON PAGE 14)**

16A. In your current or most recent SCIENCE class you took within the last two years, how often did you...

19. Have you taken a mathematics class in the last two years?

Yes . . . . . . . . 1 **(GO TO Q.19A)**
No . . . . . . . . 2 **(SKIP TO Q.22 ON PAGE 16)**

19A. In your current or most recent MATH class you took within the last two years, how often did you...

22. Have you taken a vocational class in the last two years?

Yes . . . . . . . . 1 **(GO TO Q.22A)**
No . . . . . . . . 2 **(SKIP TO Q.23)**

22A. In your current or most recent VOCATIONAL class you took within the last two years, how often did you...

Question 22 poses an additional problem. Twenty-four to twenty-six percent of the 384 respondents who should have answered Q22A to Q22F (i.e., those who did not indicate that they had not taken a vocational course) left the subpart questions unanswered. Removing the confusing filter question from between the question introduction and the associated subquestions, as suggested above, may help.

However, two other factors may account for the problem. First, students may not know what kinds of courses qualify for the label "vocational". Students' confusion regarding the term was reported by field staff; as indicated in Appendix E, there were numerous student inqueries about Q22. Anecdotal data from Survey Day Report Forms identify the term "vocational" as a confusing term in this and other questions. (The term appears in questions 11, 13I, 31, 35I,44D, 45, 50A and B, 56A, 65A, 67, and 79F and I, sometimes referring to high school courses and sometimes to postsecondary education options.) Of the 84 Survey Day Report Forms completed and submitted, 11 noted a problem with this term -- most associated with Q22, some among the other items listed above. These 11 report forms indicate that about half the schools in which the confusion was serious enough to report were private/preparatory high schools unlikely to offer such a program; the remaining half were public comprehensive high schools. We recognize that the use of the term "vocational" has generated lengthy discussion at previous questionnaire development meetings. If cross-wave consistency is of paramont importance, we recommend that various options -- such as including examples of vocational courses at Q22 and using Q65A as a list of examples to clarify students' thinking when interviewers are asking to define the term -- be tested in cognitive inteviews to identify the optimal solution to this problem.

Second, some of the subquestions are wordy and abstract. Third, estimating the degree of emphasis placed on the various activities may be a difficult judgment to make. Because of the various problems with this item, we suggest that the item be subjected to cognitive interviewing to obtain suggestions on how the item may be made more concrete, explicit, and understandable.

**Analysis: Q14 Talent Search and Upward Bound are programs that help economically disadvantaged high school students to prepare for entering and succeeding in college. At any time during high school have you participated in these programs or a similar program?**

Relatively high nonresponse was found for item 14B, the dependent part of a filter/dependent question set in which the respondent is first asked whether he or she participated in Talent Search or Upward Bound programs. In Q14B the respondent is asked to "Circle all grades during which you participated in this program".

There are two aspects of this question that may make it confusing to the respondent. The first is a minor point. The question stem asks for "grade" while the response categories are presented in terms of the student's class (Freshman, Sophomore, Junior, Senior). To be consistent with the stem, the response categories should say "ninth", "tenth", "eleventh", and "twelfth", or the stem should be changed to ask for "year".

Another inconsistency, this time between the filter and the dependent question, may seriously confuse respondents. In the filter question the respondent is asked to indicate his or her participation in several programs by circling all of the types of programs in which he or she has participated. However, in the text to the dependent question the reference is to a single program. This change for "programs" in Q14 to "program" in Q14A may cause some confusion about how to answer Q14A.

**Recommendations:** We recommend the following format for this question.

14. Talent Search and Upward Bound are programs that help economically disadvantaged high school students to prepare for entering and succeeding in college. At any time during high school have you participated in these programs or a similar program?

No . . . . . . . . . . . . . 1 (Skip to Q.15)

Yes . . . . . . . . . . . . . 2 (Go to Q.14A)

14A. Please circle the years in which you participated in Talent Search, Upward Bound, or a similar program. **(CIRCLE ALL THAT APPLY ON EACH LINE)**

| | Freshman | Sophomore | Junior | Senior | Did not Participate |
|---|---|---|---|---|---|
| Talent Search | 1 | 2 | 3 | 4 | 8 |
| Upward Bound | 1 | 2 | 3 | 4 | 8 |
| Other, similar program | 1 | 2 | 3 | 4 | 8 |

**Questionnaire Section 3: Your Plans for the Future**

**Analysis:** Q49 through Q55 Postsecondary Plans

Question 48 serves as the filter for Q49 through Q55. Question 48 inquires about the student's plans to attend postsecondary school immediately after high school. Only those respondents who do not intend to go on to school right after high school are eligible to answer Q49 through Q55. Each of these dependent questions has a relatively high nonresponse rate. Generally, this nonresponse arises from 40 respondents who did not answer the filter question and did not answer most, if not all, of the dependent questions. These students were not consistent nonresponders; more than half of them answered questions immediately before and after the filter and dependent questions. In addition, more than half of these respondents reported in Q45 that they thought they would attend postsecondary school; answering Q48 should have presented no difficulty. Though nonresponse to Q48 is low -- the adjusted nonresponse rate is five percent -- the impact of this on the apparent nonresponse rates for the dependent questions warrants further investigation into possible sources and revision of the question to improve response.

Students may have failed to answer Q48 for a number of reasons. The item contains instructions for the first major skip in the instrument. Prior routing instructions are simpler, guiding respondents to skip only one question. The 40 nonresponders who also failed to answer the dependent questions may have been confused by the skip instructions and simply not answered the filter question; they then may not have re-entered the questionnaire until Q56 or Q56A. It is also likely that at least some of the respondents correctly followed the skip instructions for response category 4, but neglected to circle the number. To remedy this mechanical problem, we suggest altering the skip instruction to read "Go to Q49 after circling this choice."

The stem of Question 48 is too dense: "Do you plan to attend either a four year college/university, two year community/junior college, or a vocational, technical or trade school after high school?" The First Follow-Up main study version was much simpler -- "Do you plan to go on to college after high school?" The question could be simplified to: "Do you plan to go to a 4-year college, community or junior college, or vocational, technical, or trade school after high school?" "Vocational, technical, or trade school" could be further simplified by "some other type of school," since team leader reports indicate that twelfth grade students do not know the meaning of "vocational," and Q48 does not seek to identify the type of postsecondary school to be attended. The question then would read: "Do you plan to go to a 4-year college, community or junior college, or some other type of school after high school?"

The response options are also not in logical order, and their ordering may be a source of some nonresponse. By the spring of their senior year, most twelfth graders should know whether and when they plan to attend postsecondary school. Consequently, "Don't know" should be the terminal response option. Since 83.1 percent of respondents reported that they planned to go on to postsecondary school immediately after high school, this response category should be first, followed by "Yes, after staying out of school for one year" and "Yes, after staying out of school for over a year." "No, I don't plan to continue my education" should be the fourth response option. Some of the nonresponse to Q48 may have arisen from embedding the "Yes, right after high school" response, and its skip instruction, in the middle of the response categories.

**Analysis: Q49 Which of the following are reasons why you have decided NOT to continue your education right after high school?**

See Section 3.4, Student Questionnaire Experiments for a full analysis of nonresponse to this item.

**Analysis: Q50A Would you say that high school has prepared you for the jobs you expect to have over the next 5 years?**

The adjusted nonresponse rate for Q50A is 25 percent. Of the 49 students who did not answer this question, 40 also did not answer the filter question. When these cases are excluded, the nonresponse rate is quite low (6 percent).

**Recommendation:** Retain item as is.

**Analysis: Q50B Do you feel that you have enough skills right now for the work career that you see yourself holding 5 years from now?**

The adjusted nonresponse rate for Q50B is 28 percent. Of the 54 students who did not answer this question, 40 also did not answer the filter question. When these cases are excluded, the nonresponse rate is low (7 percent).

**Recommendation:** Change "work career" to "job or career." Otherwise, retain item as is.

**Analysis: Q51 Do you have a steady job lined up after high school?**

Question 51OV is a variable constructed from Q51, which has a circle-all-that-apply format. If a respondent circled any of the four response categories, he/she was assigned a value of .00.

Respondents who did not answer this question but were eligible, based on their response to Q48, were assigned a value of 1.00.

The adjusted nonresponse rate for Q51OV is 25 percent. Of the 49 eligible respondents who did not answer this question, 39 also did not answer the filter question. When these cases are excluded, the adjusted nonresponse rate is low (5 percent).

**Recommendation:** Retain item as is.

**Analysis: Q52 Please list up to 3 jobs that you could get after you leave high school.**

The adjusted nonresponse rate for Q52 is 40 percent. Of the 77 respondents who did not answer this question, 40 also did not answer Q48, the filter question. When these cases are excluded, the unadjusted nonresponse rate is still relatively high (25 percent). Some of the nonresponse is probably due to the question being open-ended. Also, twelfth grade students who do not intend to go on to postsecondary school immediately after high school may not, even in the second half of their senior year, have considered possible work options.

**Recommendation:** Offer a "Don't know" response option.

**Analysis: Q53 About how much money, hourly, do you expect to earn on your first steady job after high school?**

The adjusted nonresponse rate for Q53 is 31 percent. Of the 59 students who did not answer this question, 40 also did not answer the filter question. Once again, when these cases are excluded, the nonresponse rate is low (10 percent).

**Recommendation:** Retain item as is, adjusting the lowest category to reflect a new minimum wage (if necessary).

**Analysis: Q54 If you plan to join the Armed Forces, which branch do you intend to join?**

The adjusted nonresponse rate for Q54 is 24 percent. Of the 49 students who did not answer this question, 38 also did not answer the filter question. When these cases are excluded, the nonresponse rate is low (6 percent).

**Recommendation:** Retain item as is.

**Analysis: Q55 What are your reasons for joining the Armed Forces?**

Q55OV is a variable constructed from Q55, which has a circle-all-that-apply format. If a respondent circled any of the five response categories, he/she was assigned a value of .00. Respondents who did not answer this question but were eligible, based on their response to Q54, were assigned a value of 1.00.

The adjusted nonresponse rate for Q51OV is 25 percent. Of the 48 eligible respondents who did not answer this question, 38 also did not answer the Q48. When these cases are excluded, the adjusted nonresponse rate is an acceptable 11 percent.

**Recommendation:** Retain item as is.

| **Analysis:** | Q61 | Postsecondary schools to which respondent has applied |
|---|---|---|
| | Q61C | Financial aid awarded at school 1? |
| | Q61D | Accepted by school 2? |
| | Q61E | Applied for financial aid at school 2? |
| | Q61F | Awarded financial aid at school 2? |

Question 61 inquires about the two postsecondary schools to which the respondent has applied that he/she is most likely to attend. The adjusted nonresponse rates for items Q61C, Q61D, Q61E, and Q61F are 22 percent, 37 percent, 42 percent, and 46 percent, respectively. When these four items are crosstabulated with item Q61WI (the number of schools to which the respondent has applied), the actual nonresponse rates prove to be rather low.

Table 3.2.1

Crosstabulation: Q61WI QUESTION 61 WRITE IN BLANK
By Q61C AWARDED FINANCIAL AID AT SCHOOL 1?

| **Q61C-->** | **Count**<br>**Row Pct**<br>**Col Pct** | **YES** 1 | **NO** 2 | **DON'T KNOW** 3 | **NON-RESPONSE** 9 | **Row Total** |
|---|---|---|---|---|---|---|
| Q61WI | 0 | | | 1<br>.4<br>.3 | 279<br>99.6<br>59.0 | 280<br>28.1 |
| | 1 | 4<br>2.4<br>13.8 | 68<br>40.7<br>34.5 | 68<br>40.7<br>22.9 | 27<br>16.2<br>5.7 | 167<br>16.8 |
| | 2 - 23 | 25<br>5.6<br>86.2 | 128<br>28.7<br>65.0 | 223<br>50.0<br>75.1 | 70<br>15.7<br>14.8 | 446<br>44.8 |
| Missing | | | 1<br>1.0<br>.5 | 5<br>4.9<br>1.7 | 97<br>94.2<br>20.5 | 103<br>10.3 |
| | Column Total | 29<br>2.9 | 197<br>19.8 | 297<br>29.8 | 473<br>47.5 | 996<br>100.0 |

Source: NELS:88 Second Follow-Up Field Test Data.

The majority of students (59.0 percent) who did not answer item Q61C reported that they had not applied to any postsecondary schools. Only 15.8 percent of the 613 students to whom this question applied (according to their response to item Q61WI) did not answer Q61C.

Table 3.2.2

Crosstabulation: Q61WI QUESTION 61 WRITE IN BLANK
By Q61D ACCEPTED TO SCHOOL 2?

| Q61D--> Count<br>Row Pct<br>Col Pct | YES 1 | NO 2 | DON'T KNOW 3 | NON-RESPONSE 9 | Row Total |
|---|---|---|---|---|---|
| Q61WI 0 | | | 1<br>.4<br>.4 | 279<br>99.6<br>48.5 | 280<br>28.1 |
| 1 | 2<br>1.2<br>1.3 | 1<br>.6<br>4.8 | 4<br>2.4<br>1.7 | 160<br>95.8<br>27.8 | 167<br>16.8 |
| 2 - 23 | 155<br>34.8<br>97.5 | 20<br>4.5<br>95.2 | 236<br>52.9<br>97.9 | 35<br>7.8<br>6.1 | 446<br>44.8 |
| Missing | 2<br>1.9<br>1.3 | | | 101<br>98.1<br>17.6 | 103<br>10.3 |
| Column Total | 159<br>16.0 | 21<br>2.1 | 241<br>24.2 | 575<br>57.7 | 996<br>100.0 |

Source: NELS:88 Second Follow-Up Field Test Data

The majority of students (77.7 percent) who did not answer item Q61D reported that they had applied to only one or no postsecondary schools. Only 7.8 percent of those students to whom this question applied (according to their response to item Q61WI) did not answer Q61D.

Table 3.2.3

Crosstabulation: Q61WI QUESTION 61 WRITE IN BLANK
By Q61E APPLY FOR FIN AID AT SCHOOL 2?

| Q61E--> | **Count**<br>**Row Pct**<br>**Col Pct** | **YES** 1 | **NO** 2 | **NON-RESPONSE** 9 | **Row Total** |
|---|---|---|---|---|---|
| Q61WI | 0 | | 1<br>.4<br>.5 | 279<br>99.6<br>45.8 | 280<br>28.1 |
| | 1 | 1<br>.6<br>.5 | 6<br>3.6<br>3.1 | 160<br>95.8<br>26.3 | 167<br>16.8 |
| | 2 - 23 | 195<br>43.7<br>99.5 | 183<br>41.0<br>95.8 | 68<br>15.2<br>11.2 | 446<br>44.8 |
| | Missing | | 1<br>1.0<br>.5 | 102<br>99.0<br>16.7 | 103<br>10.3 |
| | Column Total | 196<br>19.7 | 191<br>19.2 | 609<br>61.1 | 996<br>100.0 |

Source: NELS:88 Second Follow-UP Field Test Data

The majority of students (72.1 percent) who did not answer item Q61E reported that they had applied to only one or no postsecondary schools. Only 15.2 percent of the students to whom this question applied (according to their response to item Q61WI) did not answer Q61E.

Table 3.2.4

Crosstabulation: Q61WI QUESTION 61 WRITE IN BLANK
By Q61F AWARDED FINANCIAL AID AT SCHOOL 2?

| **Q61F-->** | **Count<br>Row Pct<br>Col Pct** | **YES 1** | **NO 2** | **DON'T KNOW 3** | **NON-RESPONSE 9** | **Row Total** |
|---|---|---|---|---|---|---|
| Q61WI | 0 | | | 1<br>.4<br>.5 | 279<br>99.6<br>44.1 | 280<br>28.1 |
| | 1 | | 5<br>3.0<br>3.8 | 2<br>1.2<br>.9 | 160<br>95.8<br>25.3 | 167<br>16.8 |
| | 2 - 23 | 18<br>4.0<br>100.0 | 124<br>27.8<br>95.4 | 212<br>47.5<br>98.1 | 68<br>15.2<br>11.2 | 446<br>44.8 |
| | Missing | | 1<br>1.0<br>.8 | 1<br>1.0<br>.5 | 101<br>98.1<br>16.0 | 103<br>10.3 |
| | Column Total | 18<br>1.8 | 130<br>13.1 | 216<br>21.7 | 632<br>63.5 | 996<br>100.0 |

Source: NELS:88 Second Follow-Up Field Test Data

The majority of students (69.4 percent) who did not answer item Q61E reported that they had applied to only one or no postsecondary schools. Of the 446 students to whom this item was applicable (according to their response to Q61WI), 92 (or 20.6 percent) did not answer the item.

**Recommendation:** Though actual nonresponse rates for items Q61C, Q61D, Q61E, and Q61F are relatively low, response to the items could be improved by reformatting Q61. In its present form, the question is not well-structured; items are not well-labeled; and respondents are not given explicit routing instructions. A suggested reformatting follows.

61. Have you applied to any schools?

Yes ......................... 1 (GO TO QUESTION 61A)

No ......................... 2 (SKIP TO QUESTION 62)

61A. To how many schools have you applied? |__|__| (WRITE IN)

Write below the names and addresses of the two schools to which you've applied that you're most likely to attend.

SCHOOL 1 ............. |__|__|__|__|__|__|__| FOR OFFICE USE ONLY

____________________________ ____________________________

Name of School City State

| | | |
|---|---|---|
| a. Were you accepted? | b. Did you apply for financial aid? | c. Were you awarded financial aid? |
| Yes ......... 1 | Yes .... 1 (GO TO c.) | Yes ............. 1 |
| No .......... 2 | No ..... 2 (SKIP TO SCHOOL 2) | No .............. 2 |
| Don't Know ... 3 | | Don't know ....... 3 |

If you applied to one school, skip to Question 62.

SCHOOL 2 ............... |__|__|__|__|__|__|__| FOR OFFICE USE ONLY

____________________________ ____________________________

Name of School City State

| | | |
|---|---|---|
| a. Were you accepted? | b. Did you apply for financial aid? | c. Were you awarded financial aid? |
| Yes ......... 1 | Yes .... 1 (GO TO c.) | Yes ............. 1 (GO TO Q62) |
| No .......... 2 | No ..... 2 (SKIP TO Q.62) | No .............. 2 (GO TO Q62) |
| Don't Know ... 3 | | Don't Know ....... 3 (GO TO Q62) |

**Analysis:** **Q66A Job respondent expects to hold after high school**
**Q66B Job respondent expects to hold at age 30**

The adjusted nonresponse rates for items Q66A and Q66B are 29 percent and 20 percent. The high nonresponse to these items may in part be due to the format of the entire question, principally the close juxtaposition of the open-ended items to Q66C and Q66D, which are in categorical form. Response to items Q66C and Q66D (category that describes the job the respondent expects to have after high school and at age 30) is considerably higher than response to items Q66A and Q66B. Respondents may skip the open-ended items and answer only the categorical items. In addition, respondents who plan to attend postsecondary school immediately after high school may skip Q66A because they consider it not applicable to them.

**Recommendation:** Question 66 was altered for the 2FU Field Test from its entirely close-ended format. The open-ended portion was added for the field test in order to refine the closed categories for the 2FU Main Study. We recommend that Q66 be divided into two questions the first asking explicitly for the job anticipated right after high school, with an additional response category "Will be in school". The second question would then elicit the job anticipated at age 30. See Section 3.9 for our recommendations on the refinement of the job categories.

**Questionnaire Section 4: Language Use**

**Analysis:** **Q75 Was the special help in the form of...**

Table 3.2.1 indicates that the nonresponse rate for Q75 is 38 percent. Upon closer examination, it appears that nonresponse for this question may not be as large a problem as indicated. Question 75 is a dependent question for filter question 74. Based on their responses to Q74, a total of 16 respondents should have answered Q75. Fifteen respondents (94 percent) of those who were supposed to answer Q75 actually did.

**Recommendation:** Retain item as is. We believe that nonresponse will not be a problem for this item in the main study.

**Analysis:** **Q76 To what degree do you think your English skills have improved in the following areas as a result of these special classes or activities?**

The adjusted nonresponse rate for Q76 indicates that 34 percent of all respondents who should have answered this question did not. It is felt that the wording of this question may be confusing to respondents, especially considering the fact that the native language of those answering Q76 is not English. An alternate wording for Q76 is proposed below.

> How much have your English skills improved in the following areas because you have participated in special classes or activities?

It should also be noted that Q74 erroneously directed some respondents to Q76. This erroneous skip instruction may account for some of the high nonresponse on Q76. The skip instruction for Q74 will be corrected for the main study.

**Questionnaire Section 5:** Your Opinions About Yourself and Your Attitudes

**Analysis:** Q94, Q95, Q96

In Q94 through Q96, all respondents who indicated in Q93 that they had a child were asked to provide information about their child(ren). The nonresponse rates for Q94 through Q96 presented in table 3.2.1 are extremely high, ranging from 83 to 100 percent. It is speculated that these high rates are due to respondents not getting to these questions and the questions being inapplicable, rather than respondents reaching these items and failing to respond.

A total of ten respondents indicated in Q93 they had children of their own. Of those ten respondents, eight (80 percent) answered the first part of Q94 (birth date for first child) and seven (70 percent) answered Q95. None of the eight respondents provided a birth date for a second child. Since the respondents completing the student questionnaire are typically 17 to 18 year olds, it is speculated that none of the respondents with children had more than one child.

With the exception of part 2 of Q94 (birth date of second child), item nonresponse for Q94 through Q96 appears to be much less of a problem than is indicated by table 3.2.1. Even when the base is limited to respondents with children, item nonresponse rates for these questions are still quite high (20 percent for part 1 of Q94 and 30 percent for Q95 and Q96). However, it is thought that this is due in large part to the fact that data for these questions are only available for ten respondents.

**Recommendation:** Nonresponse on these items is not expected to be a problem in the main study. Because no one provided a birth date for a second child, it is recommended that the main study Student Questionnaire only request the birth date of the first child. Data from the HS&B First Follow-Up support this recommendation. Of the HS&B and 1FU dropouts (those respondents most likely to have children), sixteen percent had one child, one percent of the dropout sample had two children, and three-tenths of one percent had three children.

**Questionnaire Section 6:** Money and Work

After adjusting by removing respondents who did not answer any items in this section of the questionnaire, no items were found to exceed the 20 percent nonresponse criterion.

**Questionnaire Section 7:** Your Family

Section 7 presented special difficulty for the determination of item nonresponse. Many respondents either did not answer any of the questions in the section at all, or were unable to finish the section, exiting the sequence at various stages. Thus, many items in the section had nonresponse greater than the 20 percent cutoff used for the other sections. Because of the sharp decline in participation throughout this section, the adjustment applied to questionnaire sections 4, 5, and 6 was not adequate. A different type of adjustment, based on the relationship between item position and nonresponse was applied.

Linear regression was used to estimate the sequential exit pattern by modelling the exit rate as a function of item position. First, percent nonresponse was regressed on item position for items in section 7. The unstandardized regression coefficient was .21, indicating that, on average, about two-tenths of a percent of respondents who made it to section 7 dropped out on each successive item within the section. Next, standardized residuals were computed. Finally, items with standardized residuals greater than 1.0 (one standard deviation greater than the predicted nonresponse value) were selected for assessment below. Only three items had a standardized residual score of greater than 1.00. These items were: Date of Completion, Q111, and Q112.

**Analysis:** Date of Completion

This question had a 46 percent nonresponse rate. The high nonresponse rate for the date of completion may be due to its placement as the very last item.

**Recommendation:** We suggest that it be put at the front of the questionnaire and numbered so that it is perceived as a question.

**Analysis: Q111 On average, how many hours per day are you responsible for their care?**

This question had a 57 percent nonresponse rate, substantially higher than predicted by its position. The question refers to the respondent's own child, younger brothers or sisters, and other relatives from Q110.

The wording of this question is difficult. "On the average" is a term often used colloquially. However, respondents may have difficulty applying the concept to this situation, especially when the time engaged may vary considerably across days.

**Recommendation:** We suggest a simpler wording, such as "About how many hours each day are you responsible for their care? ".

**Analysis: Q112 In a typical month, how many school days do you miss because of taking care of your own child or your brothers and sisters?**

This question had a 46 percent nonresponse rate, substantially higher than predicted by its position. Like Q111, the question refers to younger brothers, sisters or other relatives from Q110.

There are a number of problems with this question. First, the frame of reference is changed from that established in Q110, that is, the reference to other relatives is dropped in Q112. This may confuse respondents. Second, the item is cast in the present tense and in a hypothetical form. This may also confuse respondents. Third, the term "typical" is vague.

**Recommendations:** We suggest that the question be made consistent with Q110 and that the wording be changed to be more concrete and specific. For example, the question could be worded as follows: "Since the beginning of the school year, how many school days did you miss because of taking care of your own child, your younger brothers or sisters, or other relatives?" With the date of administration on the data file, researchers will be able to estimate a monthly rate if they so desire.

### 3.3 Logical Consistency of Responses to Filter and Dependent Questions

Questionnaire skip instructions are intended to route respondents through the questionnaire in the most efficient and logical manner by directing respondents away from questions that do not apply to them. While skip patterns can improve data quality and decrease respondent burden, if not followed correctly, they can also lead to data inconsistencies.

The Student Questionnaire contains a number of filter questions which are intended to skip respondents over specific questionnaire items, and in some cases, to skip respondents over large sections of the questionnaire. This section contains an analysis of the logical consistency of responses to filter and dependent questions. Only filter questions that were problematic are discussed here. If a filter question is not discussed in this section, it can be assumed that it was found to work satisfactorily. Table 3.3.1 below lists all filter and dependent questions from the field test Student Questionnaire.

**Table 3.3.1 - Student Questionnaire Filter and Dependent Questions**

| Filter Question | Dependent Question(s) |
|---|---|
| Q9 | Q10, Q10A |
| Q14 | Q14B |
| Q16 filter | Q16 to Q18 |
| Q19 filter | Q19 to Q21 |
| Q22 filter | Q22 |
| Q28 | Q29 |
| Q41 | Q42 |
| Q48 | Q49 to Q55 |
| Q54 | Q55 |
| Q55 | Q56 to Q65 |
| Q56 | Q57 to Q65 |
| Q61 (number of schools applied to) | Q61A-F |
| Q70 | Q71 to Q81 |
| Q93 | Q94 to Q96 |
| Q102 | Q103 to Q108 |
| Q110 | Q111, Q112 |

**Analysis:** Filter question: Q9
Dependent questions: Q10, Q10a

In Q9, respondents were asked to indicate when they were last absent from high school. Those who said they never had an unexcused absence from high school were instructed to skip to Q11, and all others were instructed to continue with Q10 and Q10a. A total of 443 respondents claimed to have never had an unexcused absence from high school. Of those, 15 respondents (3 percent) failed to skip Q10, and 58 respondents (13 percent) failed to skip Q10a. The skip failure rate of 13 percent for Q10a is quite high. One reason respondents may have had trouble with this series of questions is that they are spread across three pages in the questionnaire. It is possible that in moving from page to page respondents lost track of to which question they were supposed to skip.

**Recommendation:** For the main study, it is suggested the questions be placed on a single page or on two facing pages, if possible. If this is not possible, it is recommended that the skip instructions explicitly state to which pages, respondents are to skip (e.g. SKIP TO Q.11 ON PAGE 9).

Table 3.3.1

Crosstabulation: Q10 # OF DAYS MISSED DURING LAST ABSENCE By
Q9 LAST UNEXCUSED ABSENCE FROM SCHOOL

| Q9--> Count Row Pct Col Pct | SKIP RESPONSE 1 | OTHER RESPONSE 2 | MISSING 9 | Row Total |
|---|---|---|---|---|
| Q10 | 15 | 492 | 5 | 512 |
| 0 | 2.9 | 96.1 | 1.0 | 51.4 |
| ANSWERED | 3.4 | 91.3 | 35.7 | |
| 1 | 428 | 47 | 9 | 484 |
| SKIPPED | 88.4 | 9.7 | 1.9 | 48.6 |
| | 96.6 | 8.7 | 64.3 | |
| Column Total | 443 | 539 | 14 | 996 |
| | 44.5 | 54.1 | 1.4 | 100.0 |

Number of Missing Observations = 0

Table 3.3.2

Crosstabulation: Q10A MAIN REASON FOR LAST ABSENCE By
Q9 LAST UNEXCUSED ABSENCE FROM SCHOOL

| Q9--> Count Row Pct Col Pct | SKIP RESPONSE 1 | OTHER RESPONSE 2 | MISSING 9 | Row Total |
|---|---|---|---|---|
| Q10A | 58 | 527 | 6 | 591 |
| 0 | 9.8 | 89.2 | 1.0 | 59.3 |
| ANSWERED | 13.1 | 97.8 | 42.9 | |
| 1 | 385 | 12 | 8 | 405 |
| SKIPPED | 95.1 | 3.0 | 2.0 | 40.7 |
| | 86.9 | 2.2 | 57.1 | |
| Column Total | 443 | 539 | 14 | 996 |
| | 44.5 | 54.1 | 1.4 | 100.0 |

Number of Missing Observations = 0
Source for crosstabulations above: NELS:88 Second Follow-Up Field Test Data.

**Analysis:** Filter question: Q14
Dependent question: Q14b

Question 14 asks respondents to indicate if they had ever participated in the Talent Search or Upward Bound programs, or in some other type of program for economically disadvantaged high school students. Those who said they had never participated in such a program were instructed to skip to Q15. Those who said they had participated in one of the programs were to continue with Q14b. Nearly all respondents who were instructed to skip to Q15 did so. However, ten percent of all respondents failed to answer Q14. This is felt to be an unacceptable level of nonresponse for a filter question. A discussion of this problem, along with recommendations for restructuring Q14 can be found in the item nonresponse section (section 3.2).

**Analysis:** Filter questions: Q16, Q19, Q22

In Q16 through Q22, respondents were asked about the most recent science course (Q16 to Q18), mathematics course (Q19 to Q21), and vocational course (Q22) they had taken. For each subject area, respondents were asked to indicate if they had not taken a course in the subject in the past two years and to skip to another question. Three percent of all respondents who indicated in Q19 that they had not taken a mathematics course in the last two years answered dependent Q20 and Q21. A slightly higher skip failure rate was found for Q16. Specifically, seven percent of those who said they had not taken a science course in the last two years went on to answer Q16, Q17 and Q18. Filter Q22 appeared to work quite well. Ninety-nine percent of those saying they had not taken a vocational course in the last two years, skipped the remainder of Q22.

**Recommendation:** Filter Q16, Q19 and Q22 could be made clearer by requiring respondents to explicitly state whether they have taken each of the three types of courses during the last two years. Making the filter questions more explicit should reduce the rate of skip failure. A more detailed discussion of these questions and proposed wording changes can be found in the section on item nonresponse.

**Analysis:** Filter question: Q54
Dependent questions: Q55, Q56

Question 54 asks respondents to identify which branch of the armed services, if any, they plan to join. Those not planning to join the military were instructed to skip to Q56A and those planning to join the military were to continue with Q55 and Q56. Only 119 (71 percent) of 167 eligible respondents answered Q54. All of those who indicated in Q54 that they did not plan to join the military, correctly skipped Q55. Similarly, all 39 respondents who said they planned to join the armed forces went on to answer Q55.

For the most part, respondents correctly followed the skip instructions for Q54. However, after questionnaires were printed, it was discovered that the skip instruction for this question was incorrect. NORC team leaders asked respondents to make the correction in their booklet at the beginning of the questionnaire administration. Respondents not planning to join the Armed Forces should have skipped to Q56 rather than Q56A.

**Recommendation:** Since the skip instructions worked well, it is recommended that they simply be corrected.

**Analysis:** Filter question: Q55
Dependent questions: Q56 to Q65

All respondents who said they planned to join the military were asked in Q55 to provide a reason why they were joining the armed forces. All respondents who answered Q55 were instructed to skip to Q66. A total of 46 students provided valid responses to Q55 that would have required them to skip to Q66. In general, about one third of those who should have skipped dependent Q56 to Q65, failed to do so. The skip failure rates for these questions ranged from 43 percent for Q56 to 30 percent for Q63.

**Recommendation:** It is suggested that the skip instructions for Q55 be made clearer. A total of six pages of questions separate Q55 and Q66. Skip instructions should very clearly direct those who are supposed to skip such a large portion of the questionnaire to what point in the questionnaire they are to skip. A rewording of the skip instructions for Q55 is offered below.

Additionally questions 54 an 55 could benefit from a number of other alterations. For the most part, questions 49 through 53 deal with respondents' plans for right after high school. However, Q54 does not focus respondents on this time frame. It is suggested that the wording of the question be changed to reflect this focus in order to make Q54 consistent with the rest of the questions in the section. An additional problem with Q54 is that the question's stem is confusing because two distinct quetions are being asked as one. The format of Q55 is also problematic. Respondents are instructed to circle all categories that apply but next to each category there is an instruction directing respondents to skip to Q66. These two sets of instructions are potentially contradictory and confusing.

The following rewording and reordering of questions 54 and 55 are recommended in order to eliminate the above mentioned problems.

54. Do you plan to join the Armed Forces right after high school?

Yes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 GO TO Q54A
No . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2 SKIP TO Q56

54A. What are your reasons for joining the Armed Forces?

a. To serve my country . . . . . . . . . . . . . . . . . 1
b. I need a job . . . . . . . . . . . . . . . . . . . . . . . 1
c. To receive training for future jobs . . . . . . . 1
d. To receive money for further education . . 1
e. Other . . . . . . . . . . . . . . . . . . . . . . . . . . 1

55. Which branch of the Armed Forces do you intend to join?

(CIRCLE ONE RESPONSE BELOW AND THEN SKIP TO Q66 ON PAGE 40)

Army . . . . . . . . . . . . . . . . . . . . . . . 1--
Navy . . . . . . . . . . . . . . . . . . . . . . . 2
Air Force . . . . . . . . . . . . . . . . . . . 3
Marines . . . . . . . . . . . . . . . . . . . . 4
Coast Guard . . . . . . . . . . . . . . . . . 5
National Guard or Reserve . . . . . . . 6--

(SKIP TO Q66 ON
PAGE 40 AFTER
CIRCLING YOUR
ANSWER)

**Analysis:** Filter question: Q56
Dependent questions: Q57 to Q75

Question 56 was designed to serve two purposes. The first was to ascertain how much time respondents planning to continue their education thought they would spend working and going to school. The second purpose was to skip respondents who did not plan to continue their education past the postsecondary education questions (Q56A to Q65). While the actual number of persons who failed to follow the skip instructions for Q56 was relatively small, the percentage of the total number of respondents who should have skipped was quite large. Only 24 respondents indicated in Q56 that they did not plan to continue their education after high school. Four respondents (17 percent) did not follow the skip instruction and continued with dependent questions Q56A to Q65.

**Recommendation:** Like the skip instructions for Q55, the skip instructions for Q56 would be clearer by providing those who are supposed to skip to Q66 with the page number they are to skip to along with the question number.

**Analysis:** Filter question: Q70
Dependent questions: Q71 to Q81

In Q70, respondents were asked to specify their native language. Respondents whose native language is English were instructed to skip to Q82. All others were instructed to continue with Q71. A total of 685 respondents claimed English as their native language. Approximately 33 of these individuals continued to answer the dependent questions for an average skip failure rate of five percent.

**Recommendation:** The skip failure rate could potentially be reduced further by making the skip instructions clearer in the following ways. First, the instruction for English speakers to skip to Q82 could include a page-number reference. Similar to Q55 and Q56, in Q70 English speakers are asked to skip over three pages of the questionnaire. Clearly indicating the question and page number to which the respondent is to skip should reduce skip error. Second, the separate skip instruction next to each of the non-English languages telling respondents to go to Q71 should be presented with the use of a bracket.

**Analysis:** Filter question: Q110
Dependent questions: Q111 to Q112

In Q110, respondents were asked if they babysat any of their young relatives. Those who did babysit were asked to answer Q111 and Q112, while those who did not babysit were asked to skip to Q113. Only four percent of those who were supposed skip Q111 failed to do so. However, the skip failure rate for Q112 was 12 percent.

**Recommendation:** As with previous questions, it is recommended that the skip instructions for Q110 be made more explicit by specifying both question numbers and page numbers, reducing the skip error rate to a more acceptable level.

### 3.4 Results of Student Questionnaire Experiments

In this section we discuss the results of the student questionnaire experiments described in Section 2.1.2. The first experiment tests the effect of using precoded categories on frequency estimates. The second experiment compares the "mark all that apply" response format with a format requiring respondents to give an explicit "yes" or "no" response to a series of response categories. The third experiment addresses the issue of item order (or context) effects.

### 3.4.1 Open v. Close-ended Responses

Survey researchers have long known that respondents often make errors when reporting on the frequency with which an event has occurred, and that the form of the question used to elicit this information can influence the respondent's answer. Two broad types of questions can be administered, namely closed questions and open questions. Closed questions provide the respondent with a number of categories from which to select, whereas open questions do not. Many questions in the NELS:88 Second Follow-up Field Test Student Questionnaire involve asking students to make a judgment of frequency. The biasing effect (if any) of precoded frequency categories was assessed through the use of a split-ballot experiment in which respondents were assigned at random to an open or close-ended version of the question.

The following items require frequency judgments by the respondent. In one version of the Student Questionnaire the response categories were precoded. In a second version, the questions were left open-ended.

| | |
|---|---|
| Q7. | Report of victimization experiences in school |
| Q8. | Report of problem behaviors |
| Q10. | Number of days of school missed |
| Q24. | Time spent on homework |
| Q30. | Number of courses repeated in high school |
| Q36. | Time spent in school-sponsored extracurricular activities |
| Q37. | Time spent outside reading |
| Q39. | Hours watching TV |
| Q97. | Number of cigarettes smoked in a day |
| Q98. | Number of occasions of drinking alcoholic beverage |
| Q99. | Number of times had five or more drinks in a row |
| Q100. | Number of occasions used marijuana |
| Q100. | Number of occasions used cocaine |

Table 3.4.1 shows descriptive statistics for the uncoded version of the items. Open-ended responses were collapsed into the same categories used in the precoded version. A comparison of the distributions across categories by version was conducted to determine whether the precoded categories resulted in significant distortion of responses.

Chi-square tests of independence between category frequency and response format were conducted for each of the variables. Variables with an asterisk showed significant frequency differences depending upon the format used. The following items showed differences in frequencies across categories depending upon response format. Tables comparing the collapsed open-ended versions and the precoded versions of these variables are in Appendix B Tables 3.4.2 to 3.4.15. We undertook an examination of the residuals in each of the Appendix B table cells as a way of understanding the differences between the two formats. We use the frequencies given by collapsing the open-ended version as the standard to compare frequencies in the precoded categories. Table 3.4.16 lists items showing significant discrepancies due to response form.

**Table 3.4.1. Descriptive statistics for uncoded versions of variables**

| Variable | Mean | Std Dev | Skewness S.E | Skew | Maximum | N |
|---|---|---|---|---|---|---|
| Q7A | .57 | 1.42 | 7.13 | .11 | 20 | 485 |
| Q7B | .64 | 2.95 | 7.05 | .11 | 30 | 481 |
| Q7C | .25 | 1.24 | 10.89 | .11 | 2 | 484 |
| Q7D | .16 | .63 | 6.43 | .11 | 7 | 484 |
| Q8A* | 4.17 | 5.72 | 3.78 | .11 | 60 | 484 |
| Q8B | 1.94 | 4.32 | 4.12 | .11 | 40 | 482 |
| Q8C | 4.25 | 4.55 | 2.87 | .11 | 40 | 485 |
| Q8D* | .86 | 2.16 | 4.94 | .11 | 20 | 486 |
| Q8E | .15 | 1.26 | 16.74 | .11 | 25 | 488 |
| Q8F | .06 | .58 | 18.13 | .11 | 12 | 486 |
| Q8G | .01 | .11 | 14.96 | .11 | 2 | 487 |
| Q8H | .03 | .29 | 13.22 | .11 | 5 | 486 |
| Q8I | .02 | .29 | 19.52 | .11 | 6 | 487 |
| Q10* | 1.64 | 2.32 | 4.38 | .15 | 17 | 256 |
| Q24A1* | 4.72 | 6.85 | 3.10 | .11 | 49 | 484 |
| Q24A2* | 5.73 | 6.54 | 2.32 | .11 | 49 | 486 |
| Q24B1 | 1.66 | 2.82 | 5.76 | .13 | 30 | 337 |
| Q24B2 | 2.03 | 2.52 | 2.65 | .13 | 20 | 341 |
| Q24C1 | 1.51 | 2.48 | 4.76 | .15 | 25 | 274 |
| Q24C2* | 1.71 | 2.06 | 3.06 | .15 | 16 | 274 |
| Q24D1 | 1.92 | 3.89 | 7.00 | .11 | 45 | 469 |
| Q24D2 | 2.17 | 2.60 | 2.55 | .11 | 20 | 469 |
| Q24E1* | 1.88 | 3.79 | 6.46 | .13 | 40 | 360 |
| Q24E2* | 1.84 | 2.36 | 3.11 | .13 | 16 | 363 |
| Q24F1* | 1.98 | 4.09 | 4.43 | .11 | 30 | 461 |
| Q24F2* | 2.08 | 3.90 | 5.80 | .11 | 40 | 466 |
| Q37 | 2.76 | 4.13 | 4.13 | .11 | 40 | 481 |
| Q39* | 4.74 | 6.99 | 3.91 | .11 | 50 | 468 |
| Q30 | .54 | .98 | 2.38 | .11 | 7 | 477 |
| Q36ABA | 4.11 | 6.18 | 1.61 | .11 | 40 | 473 |
| Q36ABB | 1.41 | 3.84 | 7.65 | .11 | 50 | 459 |
| Q36ABC | 2.42 | 4.83 | 3.54 | .12 | 40 | 448 |
| Q97AB* | 1.13 | 3.68 | 4.39 | .13 | 30 | 357 |
| Q98AB* | 9.32 | 14.61 | 2.39 | .14 | 80 | 302 |
| Q99AB | .70 | 2.12 | 8.55 | .13 | 30 | 349 |
| Q100ABA | 1.99 | 6.97 | 4.80 | .13 | 50 | 339 |
| Q100ABB | .35 | 2.40 | 8.23 | .13 | 25 | 347 |
| Q101AB | .25 | 2.70 | 13.19 | .13 | 40 | 346 |

Source: NELS:88 Second Follow-Up Field Test Data.

**Table 3.4.16. Items showing significant discrepancies due to response format (open- ended vs. precoded).**

| | |
|---|---|
| Q8A | Number of times respondent was late for school |
| Q8D | Number of times respondent was in trouble for ignoring rules |
| Q10 | Number of days of school missed during last absence |
| Q24A1 | Total time spent on homework in school |
| Q24A2 | Total time spent on homework out of school |
| Q24C2 | Time spent on science homework out of school |

| | |
|---|---|
| Q24E1 | Time spent on history/social studies homework in school |
| Q24E2 | Time spent on history/social studies homework out of school |
| Q24F1 | Time spent on homework for all other subjects in school |
| Q24F2 | Time spent on homework for all other subjects out of school |
| Q36 | Time spent on school-sponsored extracurricular activities |
| Q39 | Number of hours per day respondent watches television or videos |
| Q97 | Number of cigarettes usually smoked per day |
| Q98 | Number of occasions on which respondent had alcohol to drink |

More than one-third of the items tested showed differences in response patterns due to format. The most frequent result was the tendency for respondents who received the precoded format to underutilize the null category and overutilize the adjacent, low frequency category. This occurred for Q8A, Q8D, Q10, 124A1, Q24E1, Q24F1, Q24F2, and Q97. Substantial underuse of high frequency categories was found on items Q8A, Q24A1, Q24A2, Q36, Q39, and Q98.

**Recommendations.** One plausible explanation for the first result is that respondents defined their behaviors differently depending upon response format. Respondents who engaged in a behavior at a low frequency may have, for one reason or another, discounted the low frequency as inconsequential; for example, perhaps deciding that studying for less than an hour is really not studying, and having an occasional drink or cigarette is not drinking or smoking. Under this scenario, the precoded categories may have explicitly defined low frequency behavior as consequential. This would suggest that the open-ended format may result in an small underreporting bias in the aggregate. This further suggests that the precoded format may be superior to an open-ended format. In response to the second result, increasing the number of categories at the upper end for these items may provide a remedy.

Special consideration was given to items Q8A to Q8I in response to the query from OMB about adequacy of response categories, especially at the high frequency end. Items Q8D through Q8I appear not to present a problem because less than two percent of cases fall into the highest category for each item. For items Q8A, Q8B, and Q8C, a substantially larger proportion of respondents used the highest category (12.7 percent, 5.9 percent, and 9.7 percent, respectively).

Table 3.4.17 presents the number of respondents who gave responses greater than nine for the open-ended version of Q8A, Q8B, and Q8C. In each case, the majority of these respondents could be captured with a "10 to 15 times" category. A "More than 15 times" category would help differentiate those with very high frequency from the rest of the sample. We recommend this change and addition of categories.

### 3.4.4. "Mark all that apply" and response order effects.

Often for the convenience of the respondent, a question which requires response to a number of related items, and indication of which is applicable is cast in a "mark all that apply" format. In this format, the respondent ostensibly reads down the list of items and circles a number next to the item corresponding to "yes" indicating of the item's applicability. Nonapplicability is implied by the absence of a circle.

While easing the respondent's burden, this practice has two drawbacks. First, it creates an ambiguity between a "no" response and a nonresponse which causes difficulties in data cleaning. Second, it encourages laziness in the respondent. Rather than reading each item, respondents may engage in "satisficing", or reading and responding to enough of the items to satisfy some internal criteria for adequacy of response. If this is the case, we may expect that respondents are more likely to consider items at the top of the list than at the bottom. Thus, underestimates of the applicability of items at the bottom of the list may result. A format likely to avoid this respondent tendency is one in which respondents are forced to indicate either "yes" or "no" to each item in the list. The following items were examined in this study.

**Table 3.4.17. Number and proportion of respondents giving a response greater than nine for items Q8A, Q8B, and Q8C in the open-ended version.**

Q8A # TIMES R WAS LATE FOR SCHOOL

| Value | Frequency | Percent out of 484 Responders |
|---|---|---|
| 10 | 27 | 5.6 |
| 12 | 7 | 1.4 |
| 13 | 2 | .4 |
| 14 | 1 | .2 |
| 15 | 16 | 3.3 |
| 16 | 1 | .2 |
| 17 | 1 | .2 |
| 18 | 1 | .2 |
| 20 | 6 | 1.2 |
| 24 | 1 | .2 |
| 25 | 1 | .2 |
| 32 | 1 | .2 |
| 35 | 1 | .2 |
| 40 | 1 | .2 |
| 60 | 1 | .2 |

Q8B # TIMES R CUT OR SKIPPED CLASSES

| Value | Frequency | Percent out of 482 Responders |
|---|---|---|
| 10 | 14 | 2.9 |
| 12 | 2 | .4 |
| 13 | 2 | .4 |
| 15 | 4 | .8 |
| 16 | 1 | .2 |
| 20 | 5 | 1.0 |
| 25 | 1 | .2 |
| 27 | 1 | .2 |
| 30 | 1 | .2 |
| 40 | 1 | .2 |

Q8C # TIMES R MISSED A DAY OF SCHOOL

| Value | Frequency | Percent out of 485 Responders |
|---|---|---|
| 10 | 20 | 4.1 |
| 11 | 3 | .6 |
| 12 | 3 | .6 |
| 13 | 1 | .2 |
| 14 | 2 | .4 |
| 15 | 6 | 1.2 |
| 17 | 1 | .2 |
| 18 | 1 | .2 |
| 19 | 1 | .2 |
| 20 | 4 | .8 |
| 25 | 2 | .4 |
| 29 | 1 | .2 |
| 30 | 1 | .2 |
| 40 | 1 | .2 |

Source: NELS:88 Second Follow-Up Field Test Data.

| | |
|---|---|
| Q12. | How did you get into your current high school program? |
| Q31. | Which of the following people help students in job selection? |
| Q33. | Things school provides for students who plan to go to college. |
| Q49. | Reasons you will not continue your education after high school. |
| Q113. | Things that happened to your family. |

The experiment contained two manipulations, response format and item order, resulting in four conditions. Response format varied between instructions that required respondents to "mark all that apply" and those that required respondents to mark "yes" or "no" for each item. Item order varied between the standard order (called the natural order) and a reverse order. The order manipulation was included to determine whether respondents were more likely to consider response options at the top of the list, independent of the content of the response option, under the "mark all that apply" format. Table 3.4.18 presents the proportions selecting response categories under the various format and order conditions.

**Results:** Logit analyses were used to assess the effect of response format and item order on the proportion of respondents choosing each response option. Thirteen of the 48 variables showed a significant effect of response format. In each case, the "forced choice" format elicited a higher proportion of selection (i.e., positive response). Order was significant for five variables, and in each case the closer the item was to the top of the list, the more likely it was to be selected. The interaction between format and order was significant for five variables. However, the pattern of proportions for the five variables across the different format/order combinations does not suggest clearly that respondents are less likely to choose only items at the top of the list with the "circle all that apply" format. The significant effect of format, however, suggests that respondents will choose fewer items in the "circle all that apply" format.

To assess the magnitude of underreporting in the "mark all that apply" format, variables were created that counted the number of items endorsed for each of the five questions. Table 3.4.18 shows the mean number of responses by format an order. Analyses of variance shown in Appendix B (Tables 3.4.19 through 3.4.23) confirm that for each question, significantly more items were selected under the "forced choice" format than under the "circle all that apply" format. The effect of format was significant for each of the questions. On average, just over one quarter of an item (.278) less per respondent was elicited under the "circle all that apply" response format. Across the five questions and nearly 1,000 respondents, this suggests an underreporting bias due to format on the order of about 250 applicable items. Means by format and order are shown in Table 3.4.24.

It is useful to note that this pattern of underreporting in the "circle all that apply" format was observed for both the relatively high frequency questions such as Q49 (Which of the following are reasons why you have decided NOT to continue your education right after high school?) and the low frequency questions like Q113 (In the last 2 years have any of the following happened to your family?). In general, neither response order nor the interaction between order and format significantly affected reporting. The one exception is for order which was significant for Question 12. This is probably the result of the "I don't know" category being selected more frequently when it appeared at the top of the list.

### 3.4.7 Item Order (Context) Effects.

The next study examines context effects. Context effects occur when responses to items are influenced by the text of prior items (Schuman & Presser, 1981). One position is that prior items may influence the interpretation of subsequent ones or may influence the mental material the respondent can readily bring to bear on subsequent items (Tourangeau & Rasinski, 1988). In any situation in which data will be used to examine trends (such as is intended with the NELS:88 data), changes in context may effect response -- change that may be misinterpreted as a trend.

**Table 3.4.18** **Proportion selecting response categories by response format and order**

| | | FORMAT | | | | |
|---|---|---|---|---|---|---|
| Variable | Total | Circle all that apply | | Forced Choice | | |
| | | ORDER | | ORDER | | EFFECT |
| | | Normal | Reverse | Normal | Reverse | |
| Q12A | .28 | .27 | .27 | .23 | .35 | FXO |
| Q12B | .29 | .26 | .27 | .32 | .29 | |
| Q12C | .33 | .30 | .27 | .35 | .41 | F |
| Q12D | .17 | .13 | .12 | .25 | .18 | F |
| Q12E | .25 | .24 | .27 | .27 | .24 | |
| Q12F | .18 | .13 | .19 | .12 | .27 | O |
| Q31A | .23 | .16 | .18 | .33 | .24 | F |
| Q31B | .19 | .15 | .12 | .26 | .21 | F |
| Q31C | .64 | .62 | .69 | .51 | .73 | FXO |
| Q31D | .29 | .22 | .25 | .33 | .37 | F |
| Q31E | .19 | .11 | .14 | .25 | .27 | F |
| Q31F | .20 | .21 | .19 | .19 | .22 | |
| Q33A | .80 | .80 | .73 | .85 | .83 | F |
| Q33B | .76 | .76 | .69 | .81 | .79 | F |
| Q33C | .65 | .62 | .59 | .68 | .72 | F |
| Q33D | .72 | .68 | .71 | .74 | .74 | |
| Q49A | .03 | .04 | .02 | .02 | .05 | FXO |
| Q49B | .02 | .01 | .01 | .03 | .03 | F |
| Q49C | .05 | .03 | .06 | .06 | .06 | |
| Q49D | .02 | .03 | .00 | .01 | .03 | FXO |
| Q49E | .04 | .03 | .02 | .05 | .04 | |
| Q49F | .02 | .00 | .01 | .03 | .03 | F |
| Q49G | .01 | .00 | .00 | .01 | .03 | |
| Q49H | .04 | .02 | .03 | .06 | .05 | F |
| Q49I | .00 | .00 | .00 | .00 | .00 | |
| Q49J | .01 | .00 | .01 | .02 | .02 | |
| Q49K | .01 | .00 | .00 | .01 | .01 | |
| Q49L | .02 | .00 | .02 | .02 | .02 | |
| Q113A | .12 | .12 | .11 | .14 | .11 | |
| Q113B | .06 | .06 | .07 | .05 | .04 | |
| Q113C | .04 | .03 | .03 | .04 | .06 | |
| Q113D | .05 | .05 | .03 | .07 | .03 | O |
| Q113E | .05 | .04 | .05 | .07 | .05 | |
| Q113F | .10 | .12 | .04 | .12 | .10 | FXO |
| Q113G | .03 | .02 | .02 | .03 | .04 | |
| Q113H | .02 | .02 | .00 | .03 | .03 | |
| Q113I | .01 | .00 | .01 | .02 | .02 | |
| Q113J | .01 | .01 | .01 | .02 | .01 | |
| Q113K | .26 | .22 | .29 | .24 | .29 | O |
| Q113L | .04 | .04 | .02 | .03 | .05 | |
| Q113M | .02 | .02 | .01 | .03 | .03 | F |

**Table 3.4.18 - continued**
**Proportion selecting response categories by response format and order**

| | | FORMAT | | | | |
|---|---|---|---|---|---|---|
| Variable | Total | Circle all that apply | | Forced Choice | | |
| | | ORDER | | ORDER | | EFFECT |
| | | Normal | Reverse | Normal | Reverse | |
| Q113N | .01 | .01 | .01 | .00 | .02 | |
| Q113O | .00 | .00 | .00 | .00 | .01 | |
| Q113P | .01 | .00 | .01 | .02 | .01 | |
| Q113Q | .14 | .08 | .18 | .10 | .20 | O |
| Q113R | .01 | .00 | .02 | .00 | .00 | |
| Q113S | .09 | .06 | .08 | .08 | .11 | |
| Q113T | .07 | .04 | .10 | .07 | .08 | O |

**Table 3.4.24. Mean Number of Responses by Format and Order**

| | Forced Choice | Mark All That Apply | Natural | Reverse |
|---|---|---|---|---|
| Q12 | 1.35 | 1.64[a] | 1.43 | 1.57[b] |
| Q31 | 1.52 | 1.95[a] | 1.66 | 1.80 |
| Q33 | 2.79 | 3.08[a] | 2.97 | 2.90 |
| Q49 | .17 | .35[a] | .24 | .29 |
| Q113 | 1.02 | 1.22[b] | 1.07 | 1.18 |

Note: Superscripted letters indicate a significant mean difference between conditions at less than the probability levels given below.

a) p<.01
b) p<.05

Source: NELS:88 Second Follow-Up Field Test Data.

The study examines the effect of question placement on students' general evaluations of their schools and teachers (Q6A to Q6Q). In the NELS:88 Base Year survey, student respondents made these evaluations after rating problems in the school. In the First Follow-Up questionnaire students made the general evaluations before answering questions about specific problems. Thus, changes in evaluations of schools from grade school to high school are confounded with changes in context.

While the Second Follow-Up Student Questionnaire does not contain the school problem items used in the Base Year and First Follow-Up, there are three items that record individual negative school experiences that also precede, but not directly, the general school evaluation items in the Base Year but not in the First Follow-Up. Thus the context effect study had the following design.

**Version 1:** The general-to-specific order, used in the 1FU and proposed for the 2FU, was employed. This ordering placed Q6A-6Q ("How much do you agree with each of the following statements about your current school and teachers?") before Q7A-7D ("In the first semester of the current school year, how many times did any of the following things happen to you at school?").

**Version 2:** The reverse order, asking about specific problems (Q7A-7D) before the general evaluation (Q6A-6Q), was used.

**Results:** The main objective of this experiment is to assess whether bringing to mind problems by asking respondents about negative events that might have happened to them will affect their evaluations of their school. Therefore, comparisons of average responses to the school evaluation items (Q6A-6Q) were made for the two question order groups. Comparisons of responses to Q7A-7D for the different contexts were also made. For comparability, the collapsed versions of Q7A-7D were used (see section 1). None of the results was significant, suggesting that item order had no effect on either evaluations of the school or reports of school-related behavior.

### 3.5. Item nonresponse by item position

It is useful to examine the percentage of nonresponse to items as a function of whether they appear early or late in the questionnaire. This gives some indication of whether there were too many items relative to the amount of time allotted for responding. Inferences about time are flawed because item position is confounded with item topic (grossly indicated by the topical themes of each of the seven sections of the questionnaire). With this in mind, some sense of the sheer burden imposed by the number of items can be obtained through examination and cautious interpretation of the relationship between nonresponse and item position.

Each item was recoded such that a valid response received a "0" and a nonresponse--indicated by a missing value, a "Don't know" response, or an explicit refusal -- received a "1". A "circle all that apply" item was coded as missing if none of the options were circled. For items only intended for a portion of the sample (e.g., language items to be answered by bilingual students), respondents who were supposed to skip an item or set of items were not counted in the denominator. Nonresponse was calculated using only legitimate responders as the base.

A thorough analysis of item nonresponse as it applies to individual items is given in section 3.2. In this section, we concentrate on nonresponse as a function of item position. Because of this, we eliminated certain high-nonresponse items from our analysis. These items were those we judged to be extremely problematic because of their content or format. For example, the two items asking about the birth month and year of the respondent's second child (Q94A and B) were not completed by any respondent. Similarly, high nonresponse was found for items appearing at the beginning of Q16 and Q19, asking respondents whether they had taken science (Q16) or math (Q19) in the last two years. Such winnowing of items left a total of 447 to examine. Table 3.5.1 shows the average proportion of nonresponse over all the items and by section.

Table 3.5.1 indicates that the average nonresponse across all items included in the analysis is less than 20 percent. The analysis by section shows that nonresponse increases steadily across sections. The one exception is Section 4, which shows almost as much nonresponse as Section 7. Section 4 is the language use section and is supposed to be answered only by respondents whose native languages are not English. The relatively high nonresponse across these items may indicate the difficulty of these items for respondents with English as a second language.

Table 3.5.1 also indicates that both the standard deviation in nonresponse and the number of items vary considerably across sections. A finer-grained analysis was conducted by graphing the percent nonresponding by item. This graph is shown in Figure 3.5.1. With a few exceptions, the graph shows a steady increase in nonresponse across items. The correlation between the variable's position in the questionnaire and nonresponse is .77[4]. The unstandardized regression coefficient predicting nonresponse from item position was .09140, indicating that each additional item increased nonresponse by about nine one-hundredths of one percent. Another way of thinking about this is that, assuming the linear model is a reasonable approximation of the relationship between nonresponse and item position, item nonresponse increased on average by about one percent for every 11 items in the questionnaire.

Table 3.5.1. **Descriptive nonresponse statistics across items overall and by section of the NELS:88 Second Follow-Up Field Test Student Questionnaire.**

| | Mean | Std Dev | Items |
|---|---|---|---|
| Overall | .1969 | .1642 | 447 |
| SECTION 2 | .0495 | .0646 | 116 |
| SECTION 3 | .1071 | .0956 | 119 |
| SECTION 4 | .4224 | .1042 | 42 |
| SECTION 5 | .2153 | .0353 | 100 |
| SECTION 6 | .3750 | .0203 | 14 |
| SECTION 7 | .4463 | .0751 | 56 |

Source: NELS:88 Second Follow-Up Field Test Data.

## 3.6 Scale Reliabilities

This section examines the reliability estimates for the two attitude scales in the Student Questionnaire. The Locus of Control (Rotter, 1966) and Self Esteem (Dunteman, Peng & Holt, 1974) scales have been used in the two prior rounds of NELS:88, as well as in NLS-72 and HS&B. The individual items that make up each scale are presented in Table 3.6.1.

We used Cronbach's coefficient alpha as the basic measure of reliability. Coefficient alpha is essentially a measure of the internal consistency of a scale. It is the most appropriate reliability measure for constructs that change over time, such as self-esteem or locus of control (Ghiselli, Campbell and Zedeck, 1981). In addition to computing coefficient alpha, we examined item-to-total correlations and the alpha for the remaining items if an item was deleted. Thus, one can identify

---

[4] The pattern in Figure 3.5.1 is not quite linear. A correlation between item position and the natural log of the percent nonresponding was computed. The log transformation resulted in a slightly greater correlation between position and nonresponse (r=.82). For present purposes, the linear model seems adequate.

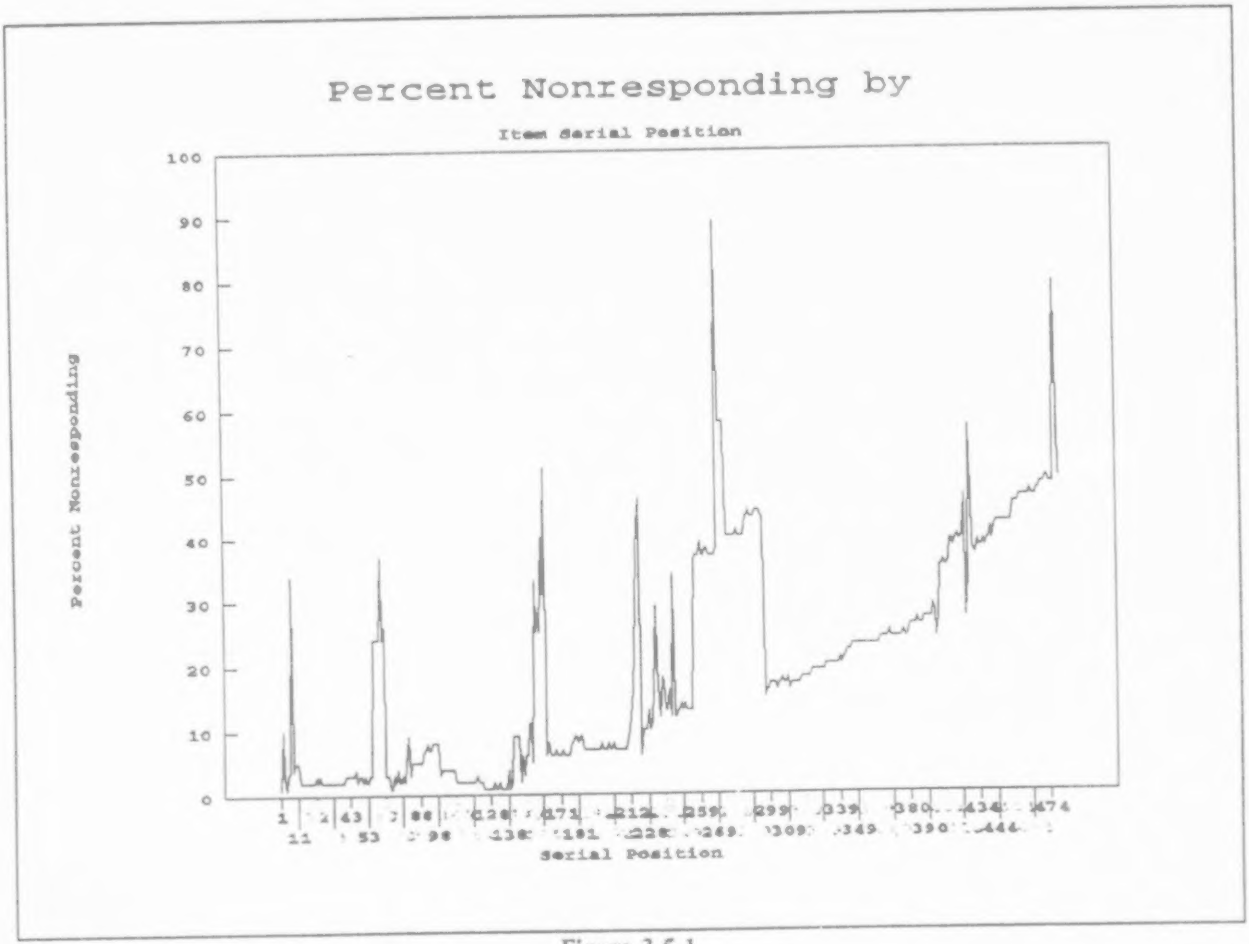


Figure 3.5.1

those items that make relatively small, or even negative, contributions to the overall reliability of the scale, as well as those that make large contributions.

**Background:** In an effort to increase the reliabilities of these two scales, the NELS: 88 Base Year Field Test modified both the locus of control and self-esteem items used in NLS-72 and HS&B. All of the items were reworded so that they could be understood by typical eighth graders. In addition, the locus of control items were given a more a personal reference than before, and the items were balanced in order to prevent response set bias. Finally, new items were added to both composites. HS&B had six items for both scales, but, to maintain comparability with NLS-72, only four of the items were actually used in calculating both of the composites. As modified for the Base Year Field Test, the locus of control composite was based on nine items and the self-esteem composite was based on ten. The reliabilities for the two composites, reported in the Base Year Field Test (Ingels et al., 1987), were .60 for the locus of control scale (vs. .56 observed in the HS&B data for the four-item scale used there) and .82 for the self-esteem scale (vs. .66 in HS&B with a four-item composite).

Based on the item-to-total correlations and alphas after the deletion of specific items, the project staff concluded that only six of the nine locus of control items and seven of the ten self-esteem items should be used in the Base Year Main Study.

**Table 3.6.1 Individual Items For Composite Variables**

| Scale | Question Number | Content |
|---|---|---|
| Self-Esteem | 82A | I feel good about myself |
| | 82D | I feel I am a person of worth, the equal of other people. |
| | 82F | I am able to do things as well as most other people. |
| | 82I | On the whole, I am satisfied with myself. |
| | 82J | I feel useless at times. |
| | 82K | At times, I think I am no good at all. |
| | 82M | I feel I do not have much to be proud of. |
| Locus of Control | 82B | I don't have enough control over the direction my life is taking. |
| | 82C | In my life, good luck is more important than hard work for success. |
| | 82G | Every time I try to get ahead, something or somebody stops me. |
| | 82H | My plans hardly ever work out, so planning only makes me unhappy. |
| | 82L | When I make plans, I am almost certain I can make them work out. |
| | 82N | Chance and luck are very important for what happens in my life. |

Source: NELS:88 Second Follow-Up Field Test Data

The First Follow-Up Field Test included the same thirteen items retained in the Base Year Main Study. Overall, the coefficient alpha for the six-item locus of control scale was .73, an improvement over the reliability of .60 for the nine-item scale tested in the Base Year Field Test. The reliability coefficient for the seven-item self-esteem scale was .82, the same as in the Base Year Field Test. Examination of the First Follow-Up Field Test results led to retaining all thirteen items in the First Follow-Up Main Study; all thirteen had reasonable item-to-total correlations (the lowest was .38) and each item increased the overall scale alpha. Table 3.6.2 summarizes the previous reliability estimates for the two scales.

Table 3.6.2

**Reliability Estimates for HS&B/NLS-72, NELS:88 Base Year Field Test, Base Year Main Study, and First Follow-Up Field Test**

| | Study | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Scale | HS&B/NLS-72 | | NELS:88 BYFT | | NELS:88 BYMS | | NELS:881FUFT | |
| | Items | α | Items | α | Items | α | Items | α |
| Locus of Control | 4 | .56 | 9 | .60 | 6 | .68 | 6 | .73 |
| Self Esteem | 4 | .66 | 10 | .82 | 7 | .79 | 7 | .82 |

Note. NELS:88 Base Year Field Test and NELS:88 First Follow-Up Field Test are based on the same sample; NELS:88 Base Year Main Study and NELS:88 First Follow-Up Field Test are based on the same items.

**Results:** The Second Follow-Up Field Test retained the same seven self-esteem items used in the two earlier rounds of NELS:88 (see Table 3.6.1 for the specific items). Further, the coefficient alpha estimated from the 805 Second Follow-Up Field Test respondents who answered all seven self-esteem questions was .83, almost identical to the estimates from the Base Year Main Study and from the First Follow-Up Field Test.

In another attempt to increase the reliability of the locus of control scale, we included an additional item ("People who accept their condition in life are happier than those who try to change things"), besides the six listed in Table 3.6.1. This item was one of the four items from which the locus of control scores were derived in HS&B and NLS-72, and its inclusion in NELS:88 would facilitate comparisons across the studies. However, the reliability of the locus of control scale is higher when this item is deleted (coefficient alpha of .71 based on 811 Second Follow-Up Field Test respondents) than it is when the item is included (coefficient alpha of .66). The decision to drop the item from NELS:88 in the Base Year thus seems to have been a sound one. This item, therefore, should be omitted from the 2FU main study locus of control items.

### 3.7 Student Data Quality

This section describes several analyses concerning the quality of the student data that do not fall under any other heading. For the most part, the analyses reported here compare the agreement between responses to closely related items in the Student Questionnaire. Other analyses examine the problems associated with specific response formats.

**Time Spent on Homework - Part vs. Whole:** Question 24, which asks about the amount of time spent on homework, is likely to play an important role in analyses of the NELS:88 data. The pretest items begin by asking about the overall time spent on homework in school (24A2) and outside of school (24A2); it continues with similar questions asking about homework in particular subjects, including mathematics (24B1 and 24B2), science (24C1 and 24C2), english (24D1 and 24D2), history and social studies (24E1 and 24E2), and all remaining subjects (24F1 and 24F2). Such item sequences, which include both an overall or general question as well as more specific items, are often used in surveys, but they are prey to some well-known difficulties. For instance, when the specific questions come first, they can change answers to the later general question (e.g., Schuman & Presser, 1981). Further, it is possible to use either open or closed items to ask these questions. The two response formats were experimentally varied in the pretest and we discuss the results of that experiment elsewhere. (See section 3.4 for results of experiments in the Student Questionnaire.) Here we focus on the consistency of the responses to the general and the specific questions; that is, we examine whether the whole is equal to the sum of the parts.

Using data from the respondents who received open versions of the homework items, we compared the total amount of time spent on homework in school as reported on the initial general item (Q24A1) with the sum computed from responses to the specific items (i.e., Q24B1, Q24C1, Q24D1, Q24E1, and Q24F1). We repeated these analyses with the items on time spent on homework outside of school (Q24A2 vs. Q24B2, Q24C2, Q24D2, and Q24F2).

In both cases, the agreement between the sum of the responses to the specific items and the reported total is less than perfect. There is exact agreement between the reported total and computed sum less than a third of the time: for the in-school items, the proportion of cases in which the sum equalled the reported total was 31.1 percent (based on 456 respondents who completed all the relevant items); for the out-of-school items, the corresponding figure was 28.8 percent (based on 459 respondents). Correlations computed between the sums and reported totals also revealed less than perfect agreement; the correlations were .627 for the in-school items and .691 for the out-of-school items. Disagreement tended to be in a specific direction -- the sum of the specific responses were larger than the amounts reported in the initial overall items. In the initial items, students reported spending an average 4.7 hours per week on homework in school and another 5.7 hours per week outside of school; these figures contrast with the averages of 7.3 and 8.0 for the sums computed from the in-school and out-of-schools items retaining the specific items, and calculating the overall from the specific components. This would retain the questions collecting data comparable to that obtained in the 1FU (although overall time spent was also included in the 1FU series).

Analysis of the data from the respondents who received closed versions of the homework item is more complicated, but points to similar conclusions. For instance, more respondents selected "None" in response to the initial overall item than selected "None" on all six of the specific items.

We suspect that the sums computed from the specific items are more accurate than the overall reports. Several considerations support this hypothesis. First, the specific items should lead to a more systematic canvass of memory than the general items and should therefore prompt fuller recall. Second, even if the more specific items have no particular value as retrieval cues, they do require respondents to continue thinking about the issue. Increased time on the task is likely to improve the quality of the answers. Finally, respondents are known to make arithmetic errors and other errors of estimation. Using the specific item (and letting the computer do the arithmetic) reduces the cognitive burden on respondents. (See Burton & Blair, 1991, for a discussion of these issues.) Based on this reasoning, we recommend dropping the overall items.

**Extracurricular Activities - Response Format:** Question 34 and Q35 are similar in content but differ in format. Question 34 asks about participation in sports and related activities like cheerleading; it uses a circle-all-that-apply format that allows a number of inconsistent combinations. Respondents can indicate, for example, both that the school did not offer a particular sport and that they played on the varsity team. Question 35 asks about extracurricular activities not related to sports (such as student government and the band); it uses a subset of the response options used in Q34 but asks respondents to indicate a single answer for each activity.

The response format used in Q34 appears to create some difficulties. First, the number of cases in which no option was selected at all is relatively high; across the nine items, the rate of missing data ranges from a low of 5.0 percent to a high of 6.8 percent (based on 996 respondents). This compares to a range of 3.5 to 4.2 percent for the items comprising Q35. Second, some respondents selected more than one of the options; across the nine items, the rate of multiple codes range from 0.5 to 3.1 percent. Even when the options selected are not, strictly speaking, logically incompatible, the presence of multiple codes is likely to complicate the analysis. For example, the two items with the highest rate of multiple responses ask about "other team sports' (34f) and "other individual sports" (34g). Multiple answers on these items are likely to reflect participation in more than one sport, but the item does not include enough detail to make this clear. Finally, these items seem to be difficult to data enter; a number of coding errors were apparent.

We recommend changing Q34 in two ways. First, we recommend that the format be changed so that respondents are required to ark "yes" or "no" for each item. This will reduce the burden on the respondents, the data entry staff, and the analysts. Second, we recommend that the items asking about other sports be dropped completely or simplified. If it is important to keep the items, we recommend using a simple yes-no format (e.g., "Did you participate in any other team sports?").

**Post-High School Job - Response Format:** Question 51 asks respondents whether they "have a steady job line up after high school" and offers four seemingly mutually exclusive response options -- "Yes, I'll continue in the job I have now," "Yes, I have a new job lined up," "No, but I am looking for a job," and "No, I haven't done anything yet to get a job." As with Q34, however, this one uses a circle-all-that-apply format and produces some odd response combinations. Altogether 145 respondents answered this question, five of them giving two responses. Four of these involved respondents who indicated both that they would continue in their present job and that they had a new job lined up. It is impossible to tell if theses respondents really meant both (that is, they intend to keep their current job and also have lined up a second job) or if they just made a mistake (for example, they may have missed the work "new" and reported their current job as the one they have lined up). Aside from the issue of how to interpret multiple answers, this item, like Q34, also fosters coding errors. We recommend that this question require a single response.

**Summer After Graduation - Response Format:** Question 64 uses a circle-all-that-apply response format to elicit information on how respondents plan to spend the summer after they graduate, if they "plan to go on to school". Some of the possible response combinations are unlikely, if not contradictory. Of the 67 respondents who indicated the "do not plan to attend or go on to school until later," seven also indicated they would "work full-time to earn money to pay for school," and 14 indicated they would work part-time to earn money for school. Similarly, of the 344 respondents who indicated they would be working full-time, 19 also indicated they would be working part-time. A total of 809 gave some response to this item; 194 of these circled more than one answer (in fact, 54 respondents circled more than two answers). Because the item does not give enough detail to clarify the situations giving rise to such combinations and because the respondent is probably in a better position than the analysts to pick the single category that best describes his or her plans, we recommend that this item be changed to ask for a single response.

**College Boards - Inconsistent Responses to Related Items:** Question 46 asks whether respondents have taken or plan to take each of several college admissions tests, including the SAT and ACT; the respondent is required to give one answer regarding each test. Question 47 asks about preparation for such tests; it is a circle-all-that-apply item that lists several possible methods of test preparation (such as private tutoring). Question 47 also includes a response option for those who have not taken and do not plan to take either the SAT or ACT; this option is listed first.

There are disagreements between the two items as displayed in the crosstabulation below.

| Classification Based on Question 47a | Classification Based on Question 46 | |
|---|---|---|
| | Took/Plan to Take | Have Not Taken/ No Plans to Take |
| Took/Plan to Take | 775 | 110 |
| Have Not Taken/No Plans to Take | 25 | 89 |

Source: NELS:88 Second Follow-Up Field Test Data.

One hundred and thirty-five respondents gave inconsistent answers to the two items. Twenty-five of these indicated in Q46 that they had taken (or planned to take) one of the two tests, but then circled the initial option in Q47 (indicating that they had not taken and did not plan to take either test); 110 displayed the opposite pattern, indicating in responds to Q47 that they had to take them in Q46. Question 47 contains an implicit skip pattern; those who never took or planned to take the tests are supposed to circle the first response option and skip the remaining items. Many of the apparent inconsistencies would probably disappear if the skipitem were reformatted as a filter question or an explicit instruction were included next to the first response option.

### 3.8. Response to OMB Queries - Student Questionnaire

In this section we report on a number of analyses requested by the Office of Management and Budget (OMB) in correspondence relating to the clearance document. OMB raised questions about each of the items or sets of items listed below. In general, the questions concerned consistency and possible redundancies among certain items, and appropriateness of certain response categories.

**Analysis:** Q8, Items A-I:

These items were fielded in an open-ended form in versions A and B of the questionnaire and in a categorical form in versions C and D. Response variance and recommendations for category construction are discussed in section 3.4.1, Open versus Close-ended Responses.

**Analysis:** Q26, Item E:

OMB requested that we determine whether additional categories, for teacher's aides and other non-peer, non-teacher school personnel are necessary for this question.

This question asks the respondent whether he/she received tutoring from someone other than his/her parents during the previous two years. Item E, "Other," with a write-in line below, captures data about tutors other than those listed in items B (a teacher at the respondent's school), C (a paid tutor outside the school), and D (a peer tutor).

Q26E TYPE OF TUTOR-SOME OTHER TYPE OF TUTOR

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| YES | 1 | 85 | 8.5 | 8.5 | 8.5 |
| LEGITIMATE SKIP | 9 | 911 | 91.5 | 91.5 | 100.0 |
| | TOTAL | 996 | 100.0 | 100.0 | |
| Valid Cases 996 Missing Cases 0 | | | | | |

Source: NELS:88 Second Follow-Up Field Test Data.

Only 85 students (8.5 percent), reported receiving tutoring from someone other than the existing categories of "a teacher at the school," "a paid tutor outside the school," or "a peer tutor." Of these 85, only 13 (1.3 percent) recorded responses on the line provided below item E. The responses appear below with the frequency of mention:

| | |
|---|---|
| a friend | 6 |
| an LPP | 2 |
| a brother | 1 |
| a study group | 1 |
| a student | 1 |
| a Stanley Kaplan SAT course | 1 |
| a "qualified tutor." | 1 |

**Recommendation:** Retain the "Other" item; do not add a separate category for teacher's aide or other type of tutor.

**Analysis:** Question 28:

This question asks the respondent whether he/she has ever taken a minimum competency or proficiency test. A "Don't know" response option was added at the recommendation of the OMB reviewer. In our response to the OMB query about this item, we suggested that we assess the validity of student responses and the utility of the "Don't know" response option. We also stated that we would determine whether respondents selecting the "Don't know" category primarily are students in schools without competency tests, by comparing student responses to principals' responses to Question 37 in the School Administrator Questionnaire. Question 37 asks whether students are required to pass a minimum competency/proficiency test in order to receive a diploma.

The "Don't know" response option was well used, by 29.1 percent of the students for whom school data were available. Student respondents who opted for the "Don't know" category were almost as likely to be enrolled in schools without a proficiency test (52.7 percent) as in schools with one (46.4 percent). For the other response categories, students' responses were generally consistent with principals' responses to Question 37.

**Recommendation:** Retain the "Don't know" response category in Question 28. The percentage of students using this category in the Field Test is comparable to the proportion of High School and Beyond sophomore cohort members opting for the category when surveyed as seniors in 1982.

Crosstabulation: Q37 DOES SCHOOL HAVE MIN COMP/PROFICIENCY TEST
by Q28 TOOK MINIMUM COMPETENCY/PROFICIENCY TEST

| **Q28--> Count Col Pct** | **NO, NOT REQUIRED 1** | **NO, NOT YET 2** | **YES, TAKEN ONCE 3** | **YES, MORE THAN ONCE 4** | **I DON'T KNOW 5** | **9** | **Row Total** |
|---|---|---|---|---|---|---|---|
| Q37<br>NO 1 | 90<br>72.6 | 10<br>30.3 | 18<br>18.6 | 16<br>20.3 | 68<br>46.6 | 5<br>21.7 | 207<br>41.2 |
| YES 2 | 34<br>27.4 | 22<br>66.7 | 76<br>78.4 | 58<br>73.4 | 77<br>52.7 | 18<br>78.3 | 285<br>56.8 |
| 9 | | 1<br>3.0 | 3<br>3.1 | 5<br>6.3 | 1<br>.7 | 10<br>2.0 | |
| Column Total | 124<br>24.7 | 33<br>6.6 | 97<br>19.3 | 79<br>15.7 | 146<br>29.1 | 23<br>4.6 | 502<br>100.0 |

Number of Missing Observations = 0

(A total of 494 student cases were excluded from the analysis due to lack of school data.)

(Approximately one-third used the "Don't know" option.) Defining "competency test" in the question stem, in order to reduce use of the "Don't know" category, is not feasible, since there is no simple gloss or definition that will universally and economically convey the meaning of the term.

**Analysis:** Comparison of Responses to Q45 and Q84A and B

For Q45, response options 03 and 04 (less than two years and two years or more of vocational, trade, or business school) have been collapsed into one category (VOC/TRADE). Similarly, categories 05 through 09 (varying completion levels of college and graduate-level education) have been collapsed into "COLL PROG," to assess the number of respondents who expect to attend vocational school or college. OMB requested that we compare students' responses to Q45 and Q84, subparts A and B, to assess the consistency of responses to parallel items.

Question 45 inquires about respondents' expectations concerning their educational goals, which are well-documented predictors of future educational behavior. Question 84 asks respondents about their chances of obtaining, fulfilling, and achieving certain life goals. Specifically, Item A of

Q84 inquires about the likelihood of the respondent's completing high school, while Item B asks about the respondent's chances that he/she will go to college.

Students' responses to Q45 are generally consistent with their responses to Q84A. Of the 33 respondents who expect only to graduate from high school (according to Q45) none reported in Q84A that they considered the chances of their graduating to be very low or low, only two responded that chances were fifty-fifty, and 20 reported that chances were high or very high. Similarly, none of the respondents who expect to attend vocational, trade, or business school responded that the likelihood of their completing high school was very low or low, seven indicated that they were fifty-fifty, and 39 reported that they were high or very high. Only two students expecting to attend college reported that the chances of their completing high school were very low or low, 13 said that they were fifty-fifty, and 709 indicated that they were high or very high.

Students' responses to Q45 are also generally consistent with their responses to Q84B, though less so than with Q84A. Of the 33 respondents who expect only to graduate from high school (according to Q45), none reported in Q84A that they considered the chances of their attending college to be very high, only one responded that chances were high, eight reported that chances were fifty-fifty, and 13 reported that chances were low or very low.

For students who expect to attend only vocational, trade, or business school, the distribution of responses to Q84B is the opposite of that expected. Only three of these respondents indicated that the likelihood of their attending college was very low, seven reported that it was low, 12 said that it was fifty-fifty, and 24 responded that it was high or very high. Students expecting to attend only vocational, trade, or business school may respond inconsistently to Q84B for several reasons. Respondents may misconstrue the intention of "college" in Q84B because of its appropriation by vocational, trade, and business schools (e.g., "beauty colleges"). Also, students who expect to enroll in a vocational, trade, or business program at a community or junior college may report in Q45 that they expect to attend a vocational school rather than a college, but not make that distinction when answering Q84B.

Finally, students who reported in Q45 that they expect to attend college responded consistently to Q84B. None reported that the chances of their attending were very low, one reported that chances were low, 27 indicated that chances were fifty-fifty, and 696 responded that chances were high or very high.

**Recommendation:** Items A and B of Q84 are helpful in setting a context for the question and provide an accurate, low burden inter-item consistency check. Therefore, these items should be retained.

**Analysis:** Q46

This question asks the respondent whether he/she has taken or plans to take college admissions and placement tests (the PSAT, SAT, ACT, and AP test). Taking these tests is an important step in preparing for and choosing a college. Item E inquires about any other admissions test the respondent has taken or plans to take. OMB suggested that we add a write-in line to this question, to determine whether a category for Achievement Tests or categories for other tests are needed based on common reporting.

A high number of respondents (205), reported that they had taken or planned to take an "other" admissions test. Only 14 respondents recorded the name of the other test on the line provided. Half of these respondents reported that they had taken or planned to take achievement tests (presumably the College Board Achievement Tests). Other tests reported by students were the CLEP test, TASP, Bryant and Stratton Evaluation, and college admissions test.

**Recommendation:** Retain the "Other Admissions Test" item; additional categories are unnecessary.

**Analysis:** Q48 and Q56 (formerly 56A):

In its initial query about these questions, OMB stated that the first response option in Q56 ("I do not plan to go on to school") is unnecessary, since skip instructions in Q48 and Q55 instruct persons who do not plan to go on to postsecondary school immediately after high school to skip Q56. We responded that this response category can be deleted, if it is assumed that the routing instructions

Crosstabulation: Q45 HOW FAR IN SCHOOL WILL YOU GET? BY
Q84A CHANCE YOU WILL GRAD FROM HIGH SCHOOL

| **Q84A-->** **Count**<br>**Row Pct**<br>**Col Pct** | **VERY LOW 1** | **LOW 2** | **FIFTY-FIFTY 3** | **HIGH 4** | **VERY HIGH 5** | **MISSING DATA 9** | **Row Total** |
|---|---|---|---|---|---|---|---|
| Q45 2<br>HS GRAD ONLY | | | 2<br>6.1<br>9.1 | 4<br>12.1<br>5.5 | 16<br>48.5<br>2.3 | 11<br>33.3<br>5.6 | 33<br>3.3 |
| 3<br>VOC/TRADE | | | 7<br>9.7<br>31.8 | 9<br>12.5<br>12.3 | 30<br>41.7<br>4.3 | 26<br>36.1<br>13.3 | 72<br>7.2 |
| 5<br>COLL PROG | 1<br>.1<br>50.0 | 1<br>.1<br>100.0 | 13<br>1.5<br>59.1 | 59<br>6.8<br>80.8 | 650<br>74.8<br>92.5 | 145<br>16.7<br>74.4 | 869<br>87.2 |
| MISSING | 1<br>4.5<br>50.0 | | | 1<br>4.5<br>1.4 | 7<br>31.8<br>1.0 | 13<br>59.1<br>6.7 | 22<br>2.2 |
| Column<br>Total | 2<br>.2 | 1<br>.1 | 22<br>2.2 | 73<br>7.3 | 703<br>70.6 | 195<br>19.6 | 996<br>100.0 |

Number of Missing Observations = 0

Crosstabulation: Q45 HOW FAR IN SCHOOL WILL YOU GET? BY
Q84B CHANCE YOU WILL GO TO COLLEGE

| Q84B--> Count<br>Row Pct<br>Col Pct | VERY LOW 1 | LOW 2 | FIFTY-FIFTY 3 | HIGH 4 | VERY HIGH 5 | MISSING DATA 9 | Row Total |
|---|---|---|---|---|---|---|---|
| Q45 2<br>HS GRAD ONLY | 8<br>24.2<br>72.7 | 5<br>15.2<br>38.5 | 8<br>24.2<br>16.3 | 1<br>3.0<br>.8 |  | 11<br>33.3<br>5.6 | 33<br>3.3 |
| 3<br>VOC/TRADE | 3<br>4.2<br>27.3 | 7<br>9.7<br>53.8 | 12<br>16.7<br>24.5 | 10<br>13.9<br>7.8 | 14<br>19.4<br>2.3 | 26<br>36.1<br>13.3 | 72<br>7.2 |
| 5<br>COLL PROG |  | 1<br>.1<br>7.7 | 27<br>3.1<br>55.1 | 117<br>13.5<br>90.7 | 579<br>66.6<br>96.7 | 145<br>16.7<br>74.4 | 869<br>87.2 |
| MISSING |  |  | 2<br>9.1<br>4.1 | 1<br>4.5<br>.8 | 6<br>27.3<br>1.0 | 13<br>59.1<br>6.7 | 22<br>2.2 |
| Column Total | 11<br>1.1 | 13<br>1.3 | 49<br>4.9 | 129<br>13.0 | 599<br>60.1 | 195<br>19.6 | 996<br>100.0 |

Number of Missing Observations = 0

**Q46E TAKEN OTHER ADMISSIONS TEST**

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| HAVEN'T THOUGH ABOUT | 1 | 298 | 29.9 | 29.9 | 29.9 |
| DON'T PLAN TO TAKE | 2 | 388 | 39.0 | 39.0 | 68.9 |
| ALREADY TAKEN | 3 | 81 | 8.1 | 8.1 | 77.0 |
| PLAN TO TAKE THIS YR 8 | 4 | 124 | 12.4 | 12.4 | 89.5 |
| MISSING DATA | 9 | 105 | 10.5 | 10.5 | 100.0 |
|  | TOTAL | 996 | 100.0 | 100.0 |  |

Valid Cases 996 Missing Cases 0

| Q46EWI QUESTION 46 WRITE IN BLANK | | | | | |
|---|---|---|---|---|---|
| **Value Label** | **Value** | **Frequency** | **Percent** | **Valid Percent** | **Cum Percent** |
| DATA PRESENT | 1 | 14 | 1.4 | 1.4 | 1.4 |
| NO DATA PRESENT | 9 | 976 | 98.0 | 98.6 | 100.0 |
| | .6 | .6 | MISSING | | |
| | TOTAL | 996 | 100.0 | 100.0 | |
| Valid Cases 996 | Missing Cases 0 | | | | |

Source: NELS:88 Second Follow-Up Field Test Data.

embedded in the filter questions will be correctly followed by all students; we maintained that the item would facilitate logical consistency editing of the data files.

This is not quite accurate. Question 56 is a filter question, and the initial response option is part of the routing instructions. Students who do not plan to attend postsecondary school and who do not intend to join the armed forces are not instructed in prior questions to skip Q56. The first response option in Q56 instructs these respondents to skip the following questions about postsecondary education and go to Q66. (Students who plan to attend postsecondary school, but not immediately after high school, are routed the same as students who intend to continue their education right after high school.) This response option therefore cannot be deleted if the present format of questions 45 and 56 is retained.

**Analysis:** Q69, Items A1, B1, C1, D1, E1, F1, G1, H1:

OMB recommended that Q69 be fielded in an open-ended form rather than in categorical form, for the construction of response categories. This question asks students to estimate the annual salaries of several occupations -- bank teller, automobile mechanic, police officer, lawyer, bus driver, nurse, secretary, and mail carrier. The purpose of Question 69 is to obtain comparisons of adolescents' perceptions of earnings of this sampling of occupations.

Based on the analysis of salary frequencies for each occupation, it was determined that the optimal response categories for this question are: 1) below $15,000; 2) $15,000 - $19,999; 3) $20,000 - $24,999; 4) $25,000 - $29,999; 5) $30,000 - $34,999; and 6) over $35,000. These response categories are sufficiently narrow to permit intercategorical comparisons (rankings) without overburdening respondents.

**Analysis:** Q118, Item K

This question inquires about the amount of time parents engaged in a number of quality activities with the respondent during the previous year. The addition of item K, "Spend time with the family at home (eating evening meals or spending an evening together)," was suggested by OMB for field test evaluation.

A high percentage of respondents (78.9 percent) reported spending time with their families at home often or sometimes, while 21.1 percent reported doing so rarely or never. Note that, because of its position in the questionnaire, nonresponse was very high for this item.

**Recommendation:** Retain item K; the response variation is significant.

**Q118K HOW OFTEN YOU/PARENTS SPEND TIME AT HOME**

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| NEVER | 1 | 41 | 4.1 | 7.8 | 7.8 |
| RARELY | 2 | 70 | 7.0 | 13.3 | 21.1 |
| SOMETIMES | 3 | 184 | 18.5 | 35.0 | 56.1 |
| OFTEN | 4 | 231 | 23.2 | 43.9 | 100.0 |
|  | . | 470 | 47.2 | MISSING |  |
|  | TOTAL | 996 | 100.0 | 100.0 |  |

Valid Cases 526 Missing Cases 470

Source: NELS:88 Second Follow-Up Field Test Data.

### 3.9 Recommendations for Student Questionnaire

In this section we summarize our findings and present suggestions for ways in which the student survey instrument may be improved. In the above sections, we identify several aspects of the NELS:88 2FU Field Test Student Questionnaire that made it a rather difficult instrument with relatively high levels of nonresponse. Though we were not able to test it in this study, there is good reason to think that item nonresponse varies by respondent characteristics. For example, for the Base Year Main Study data, a composite nonresponse variable was created summing across high nonresponse items, and differences in group propensities toward item response were analyzed. Blacks and Hispanics were most likely to have high item nonresponse. Students with lower test quartiles, and males were also more likely to be nonrespondents at the item level. (See Spencer et al., 1990.) If high nonresponse on items translates into disproportionately high nonresponse for key policy-relevant subgroups such as blacks and Hispanics, then its implications for data quality are particularly unsettling.

**Length:** It must be noted that the field version of the Student Questionnaire was longer by design than the main study instrument would be expected to be. An instrument approximately 20 percent lengthier than feasible for the main study was implemented so a greater number of items could be pretested. There are advantages and disadvantages to this approach. One disadvantage is that the items in the latter sections of the instrument are not pretested with as many observations as those in earlier sections, and those responses obtained do not reflect the responses of noncompleters -- generally the poor readers and less motivated students who are key parts of the test sample. This can be remedied through testing the high nonresponse items in cognitive interviews. A second disadvantage is that it is difficult to separate item related nonresponse from questionnaire length related nonresponse for questions positioned later in the instrument. The benefit is that far more items can be pretested, even though unevenly, and a firmer sense of optimal length is obtained (firmer certainly than the information that could be obtained from a handful of timings).

However, even with those qualifications taken into consideration, one of the clearest findings was that the field test version of the Student Questionnaire was too long to be completed during the allotted 60-minute period. The high nonresponse rates for items in latter sections of the questionnaire illustrate this quite well. The Student Questionnaire will have to be significantly shortened (by almost 30 rather than 20 percent) in order to reach the goal of data completeness.

We recommend that the Student Questionnaire be shortened to approximately 90 questions for the 2FU Main Study. (In the following discussion, "question" refers to an item, whether it seeks a single response or multiple data elements.) This assessment is based not only on the 2FU Field Test, but on evidence from prior rounds of HS&B and NELS:88. The HS&B Sophomore Cohort 1FU contained 128 questions; for non-critical items, nonresponse rates were in the 10 - 20 percent range from Q85 through the end of the instrument. The level of missing data was rarely above five percent in the first two-thirds of the questionnaire. The NELS:88 First Follow-Up Student Questionnaire contained 113 questions -- fewer than in the HS&B 1FU but the questions were more dense with data elements and the format was more difficult. Item nonresponse rates for non-critical items ranged from 10 - 15 percent from Q89 on, even though field staff took extraordinary measures to give students adequate time to respond to every question. (Giving this time beyond the official limits for more complete data drove up costs and increased school burden.) The current version of the instrument clearly needs to be cut to a more optimal length of approximately 90 questions.

**Format:** The field test analysis uncovered a number of formatting problems that run the risk of undermining data quality. We found examples of overcrowded question formats in each of the NELS:88 2FU questionnaires. In the Student Questionnaire overcrowding was the likely cause of missing data about college financial aid applications and awards. An overcrowded question format is likely to lead to missing data due to confusion, or to an increase in the response time of punctilious respondents diligently trying to work their way through a tangled web. "Avoid overcrowding" should be a cardinal rule when designing self-administered questions. The cost of extra paper and printing is balanced by the benefit of more and better quality data.

As our experiment demonstrated, the use of a "mark all that apply" format may lead to underreporting. Though one may argue that this format is less burdensome for respondents, the evidence suggests that respondents hurry through the list and fail to indicate items they would have indicated if they were forced to make an explicit affirmative or negative decision about each item. Given this result, it seems clear that the "mark all that apply" format should be avoided. Questions such as Q34, in which respondents are asked to mark all that apply in a somewhat crowded matrix format, exhibit both overcrowding and "mark all that apply" problems. The format of this question and others like it should be reconsidered.

**Question/response logic and flow:** A second persistent problem concerns the logic or flow of questions and response categories. Often, the respondent was asked to indicate that he or she lacked the necessary experience to answer the question only after he or she had already read the question. This practice wasted the time of respondents for whom the question was not relevant and disrupted the logical link between the question text and the response categories by interposing the filter indicating the question was not applicable. We recommend that this practice be avoided. Filters should come before the dependent questions, rather than being included within them.

A related problem may be termed the problem of double-barreled response categories. These are response categories that try to serve the purpose of acting as filters, while also determining which of a number of choices are applicable for one of the filter categories. A question with response categories of the following type -- "No, (reason a)", "No, (reason b)", "No, (reason c)", "Yes" -- is not recommended because all possible reasons for "No" are not listed, making the task for the respondent who wishes to say "No" a difficult one. It would be better to break these types of questions into two questions, one serving as a filter and asking only for a yes or no response, and a second in which respondents who answered "no" to the first would be asked to indicate which of several reasons were applicable. The latter two-question format avoids the ambiguity posed by the one-question format. This problem is similar to the question overcrowding issue addressed earlier.

**Difficult Judgments and/or Language:** Simplified wording has been suggested for several questions. One further example is the occupation question. While the categories are meaningful to researchers studying occupation choice, it is questionable whether they are equally meaningful to respondents. A careful scrutiny of the items has led us to recommend a clarified wording. The suggested clarifications are displayed in Table 3.9.1. Other items found to be problematic in the field test would benefit from a similar scrutiny or cognitive interview testing during the summer period of revision.

Finally, some types of judgments are too difficult to make and may result in poor quality data. For example, questions asking respondents to make judgments such as "on the average" and "in a typical month" may lead to respondents making wild guesses simply because they are not used to thinking in these terms. In general, attention to simplicity in language and logic and clarity in response format along the lines suggested in this report should facilitate the construction of a well-honed Student Questionnaire for the main study.

**Table 3.9.1 Revised Categories for Occupation Question**

FARMER, FARM MANAGER

FULL-TIME HOMEMAKER

LABORER such as construction worker, car washer, garbage collector, farm hand

MANAGER such as sales manager, office manager, school administrator, retail buyer, restaurant manager, government administrator

MILITARY such as career officer or enlisted person in the Armed Forces

OFFICE WORKER such as data entry clerk, bank teller, bookkeeper, secretary, word processor, mail carrier, ticket agent

OPERATOR of machines or tools such as meat cutter, assembler, welder, taxicab/bus/ truck driver

OWNER of a small business or restaurant, contractor

PROFESSIONAL such as accountant, registered nurse, engineer, banker, librarian, writer, social worker, actor, athlete, artist, politician, but not including school teacher

PROFESSIONAL such as minister, dentist, doctor, lawyer, scientist, college teacher

PROTECTIVE SERVICE such as police officer, fire fighter, detective, sheriff, security guard

SALES such as sales representative, advertising or insurance agent, real estate agent

SCHOOL TEACHER in elementary, junior high or high school, but not college

SERVICE WORKER such as hair stylist, practical nurse, child care worker, waiter, domestic, janitor

TECHNICAL such as computer programmer, medical or dental technician, draftsperson

TRADESPERSON such as baker, auto mechanic, machinist, house painter, plumber, phone/cable installer, carpenter

NOT PLANNING TO WORK

WILL BE IN SCHOOL

OTHER

### 3.10 Preliminary Analysis of the Cognitive Test Battery

The following sections focus on the analysis of the Cognitive Test Battery data. The objectives and structure of the cognitive tests used in the Base Year, First Follow-Up, and Second Follow-Up studies are discussed, followed by analyses of the data collected in multiple choice and free response sections of the test battery in the 1FU Field Test. In conclusion, recommendations for the main study tests are submitted.

### 3.10.1 Review of Objectives and Structure of Previous Tests

**Objectives and Background:** The purpose of the NELS:88 cognitive test battery is to provide longitudinal measures of student achievement in reading, mathematics, science, and history/citizenship/geography that can be related to student background variables and educational processes. The tests must provide accurate measurement of the status of individuals at a given point in time, as well as of their cognitive growth over time. Like the earlier longitudinal studies, the National Longitudinal Study of the High School Class of 1972 (NLS-72) and High School and Beyond (HS&B) in 1980, the NELS:88 database will be used to study factors that contribute to individual differences in achievement.

Unlike NLS-72, which sampled students during the twelfth grade year only, and HS&B, which used the same test battery to evaluate students at two points in time (grades 10 and 12) the structure of the NELS:88 study requires a flexible approach to testing. Since NELS:88 will be testing students over a span of four years, gains in student achievement levels can be expected to be considerably greater than over the two-year interval covered by HS&B. In addition, the earlier cohorts, which selected students from the tenth or twelfth grade population, did not include those students who had dropped out of school before reaching the sampled grade level. By selecting a sample of eighth graders, NELS:88 includes more of the students who will eventually drop out of school than did the earlier studies. Both of these considerations -- the four year time span during which greater growth will occur, and the greater diversity of the sampled population -- suggest that the single-test approach used for NLS-72 and HS&B would not be adequate for the current study.

Accurate measurement of individual achievement requires that each student answer test items of appropriate difficulty. Items that are much too hard for a given student provide very little information about the student's skill level; nor are items that are much too easy for the student very useful. Those test items that are slightly above or slightly below a particular student's ability level are the most valuable in pinpointing the precise standing of an individual relative to the skill being measured. There are several approaches to ensuring that a student is administered test items that are appropriate to his or her level of achievement. One is simply to give a very long test, with a wide enough range of item difficulties that at least some of the test items will be appropriate for any given student. Another, computer-adaptive testing, can measure individual achievement very accurately with a small number of items by selecting the next test question from a large, pre-calibrated item pool according to the student's correct or incorrect responses to the previous items administered. Neither of these approaches is practical given the constraints of the NELS:88 study. The limited time available for testing (85 minutes for four subject areas) does not allow for tests that could be long enough to contain the items needed for all students at all points in time. The substantial development and hardware costs of computer-administered tests ruled out their use for this study. However, ideas borrowed from both of these extremes have been drawn upon in designing the structure of the NELS:88 cognitive tests.

The NELS:88 test design incorporates both grade-level-adaptive and individually-adaptive features. The tenth and twelfth grade tests need to include some more difficult test items to keep pace with the students' gains in achievement. In addition, in the later years more difficult forms of the tests can be given to students who have done well in earlier administrations, while easier versions can be developed to more precisely target the ability levels of those students who tested in the lower ranges. Through the use of Item Response Theory and common items linking all years and forms of the tests, the scores can be put on the same scale for the purpose of cross-sectional comparisons as well as measuring gains in achievement over time.

In the NELS:88 Base Year, the same set of tests was given to all students in the sample. Adapting test forms to student achievement level was not considered for two reasons. Since the greatest amount of the differentiation among students in curriculum exposure occurs later on in the high school years, a single set of tests appropriate to the vast majority of the eighth grade students could be designed. In addition, with no prior information available on the achievement level of sampled students, there would have been no basis for assignment to one test form or another. A very important consideration in constructing the grade 8 tests was to include a range of item difficulties sufficiently wide to provide individually-reliable measurement for all students. Each subtest contained some very easy items, to minimize floor effects, and some quite difficult items, to avoid ceiling effects. The test timings and number of items for each of the test sections are as follows (and remained the same for grade 10). (For more detail about the content and performance of the grade 8 tests, refer to Psychometric Report for the NELS:88 Base Year Test Battery, April, 1991.)

| | | |
|---|---|---|
| Reading | 21 Questions | 21 Minutes |
| Mathematics | 40 Questions | 30 Minutes |
| Science | 25 Questions | 20 Minutes |
| History/Citizenship/Geography | 30 Questions | 14 Minutes |

By the time the tenth grade tests were designed, two critical pieces of information were available to guide the process: statistics on overall test results in the Base Year, and the performance of each student who participated. The potential for substantial gains in achievement from two additional years of exposure to coursework suggested factors to consider in constructing the grade 10 tests.

Although there is some diversity in curriculum in history/citizenship/geography, and even more in science, the majority of high school students continue to take courses in both of these areas through the tenth grade. The number of items that could be included in these tests was judged sufficient to cover a relatively wide range of content and difficulty. In these subject areas, the average item difficulty was raised slightly from the Base Year version, and the same form was administered to all students.

In reading and mathematics, the presence of considerable numbers of both very high and very low test scores in the Base Year indicated the need for a multi-level approach to the tenth grade tests. The reasoning is somewhat different for the two subject areas, but in both cases it is related to the idea that the most useful items in a test are those closest in difficulty to the student's ability level.

The 21 questions in the reading test suggest the presence of more diversity in content and difficulty than is actually present, since the questions are based on only five reading passages. Some of the higher-scoring students could be expected to gain considerable sophistication in their reading skills by tenth grade, so for them the easier passages would be linked with not one, but an average of four, "wasted" items. On the other hand, some of the poorer readers might find more complex passages much too demanding. As a result, if the same form were administered to all students at tenth grade, too few of the test items would be within an appropriately-targeted range of difficulty as described above. Two forms of the reading test were designed for grade 10. Those students who scored below the mean in grade 8, as well as students who were not tested in grade 8, were given the easier form two years later; those who had scored above the mean received the harder form.

Three forms of the mathematics test were constructed for tenth grade, for students in the lowest quartile, middle half, and upper quartile of the grade 8 score distribution. Students being tested for the first time received the middle-difficulty form. Again, the objective of the multi-level design was to achieve individually-reliable measurement by targeting item difficulty close to the students' achievement levels. In this case, the primary consideration was the expected diversity in growth trajectories of the students over the two-year time interval. Many students take little or no mathematics beyond eighth grade, and it is reasonable to assume that they are primarily the lower scoring students. They are unlikely to have made substantial gains in this area without additional coursework. Other students, presumably those scoring at higher levels in the Base Year, had taken courses in algebra and geometry by the time they were followed up in tenth grade, and it was necessary to extend the range of difficulty by constructing harder test forms in order to avoid ceiling effects for this group.

Care was taken for both sets of multi-level tests, reading and mathematics, to include some difficult items on the easier forms, and some easy items on the harder forms. Thus measurement error in the Base Year, or administration difficulties in the First Follow-Up, would not make a test form completely inappropriate for the student.

A similar multi-level structure is anticipated for the twelfth grade test battery. Plans for which subtests will be multi-level, and how many levels of each will be required, will await analysis of the grade 10 test results and evaluation of the need for and success of this approach.

**History of the Field Tests:** One year prior to each of the NELS:88 main survey years, field tests were conducted that included cognitive test items in the four subject areas. In 1987, prior to the main survey of the eighth grade cohort, students in grades 8, 10 and 12 were tested. In 1989, only tenth and twelfth graders participated. The objective was to develop and evaluate pools of items from which the final forms of the tests could be selected for the main survey years, so many more items needed to be field tested than would eventually be needed. Since only 85 minutes of testing time were available (the same as for the main study tests) it was not possible to test all subject areas for all students. Half of the field test sample took double-length tests in Reading and Science, the other half in Mathematics and History/Citizenship/ Geography. About two-thirds of the 1989 group had also participated in the 1987 field tests. These field test cohort students were given tests in the same two subject areas in both administrations in order obtain data necessary for evaluating item-by-item gain from grade 8 to 10, and from grade 10 to 12. All subtests were analyzed for difficulty, factor structure, and internal consistency. Items for the main study tests were selected on the basis of these considerations, potential for course-related gain, and content specifications. See the NELS:88 Base Year and First Follow-Up Field Test Reports for additional information on the first two NELS:88 field tests.

### 3.10.2 Design of 1991 and 1992 Tests

**Issues for the Grade 12 (1992) Test Battery:** The same issues must be considered in designing the grade 12 tests as were important for grade 10. The difficulty of the tests needs to keep pace with student achievement in advanced courses especially in mathematics and science, while also measuring accurately for students not taking these courses, and for dropouts. The testing design, whether multi-level or otherwise, must minimize floor and ceiling effects in order to accomplish the goal of accurate longitudinal measurement that is necessary for relating achievement gains to students' educational experiences. The content of the tests must appropriately sample from the domain of knowledge the tests claim to represent.

A new issue has been introduced that could radically alter the structure of the grade 12 test battery: the possible inclusion of free response items. As originally designed, the NELS:88 test battery was to have been composed entirely of multiple-choice items. The advantages of this approach for measuring a widely diverse sample of students in a short amount of testing time are obvious. The small amount of time required per item for multiple choice tests means that a large number of items can be administered to each student. Thus a relatively short test can still cover a wide range of difficulty and a variety of content. Extensive experience in multiple choice testing over the years has resulted in a great deal of skill in writing good items and well-established procedures for evaluating their performance. Unifactorial multiple choice tests lend themselves to the NELS:88 requirement of putting different test forms on the same scale for the purposes of cross-sectional comparisons (including possible links to NLS-72 and HS&B) and measurement of gain. Their validity has been demonstrated for a wide number of applications from prediction of college grade point average to occupational and professional assessment. Some of the perceived drawbacks of multiple choice testing can be at least partially counteracted; for example, the possibility of students correctly guessing the answers to items they do not know can be addressed by Item Response Theory.[1]

But there are other considerations in measuring student achievement for which free response testing may be advantageous. Free response format in which the student must solve a problem, write an explanation, draw a diagram, etc. -- requires that the answer come entirely from the student's own knowledge and experience. There is no possibility of a response in the choice set providing a hint of the correct answer, or conversely, of a student being cued that his or her response is not correct by not finding it as one of the choices. Multiple choice format cannot easily give detailed information about the possible misconceptions that led to an incorrect answer; nor does it allow for the possibility of a test taker coming up with a different correct answer not envisioned by the test writer. Both of these situations are possibilities in free response format.

All other things being equal, inclusion of free response items on an **experimental** basis in grade 12 would seem at first glance to be desirable on its own merits. However, as is usually the case, all other things are not equal, and the potential impact of adding free response items to the NELS:88 test battery must be carefully weighed before a decision is made. Given that the 85 minutes allotted for the cognitive tests can probably not be expanded significantly, the addition of free response items would necessarily reduce the already limited time available for the four subject tests. Either test length (and reliability) would have to be compromised to an unacceptable extent, or one or more of the subtests might have to be deleted or spiralled. Deleting a subtest would eliminate the possibility of measuring longitudinal gain in that subject area, while spiralling would severely limit subgroup gain analysis.

**Design of the 1991 Field Test:** The need to evaluate the advantages and disadvantages of free response items for the purposes of the NELS:88 survey was the driving consideration in the design of the 1991 field test. Specifications for the objectives, subject area(s) and content of free response test items to be used in the 1992 main study have not yet been established. Therefore, the field test needed to contain as many items with as much diversity as possible to allow flexibility in the choices to be made subsequently. This requirement prompted a redesign of the field test structure that had been used in the two prior years.

---

[1] IRT is a method of estimating ability level by considering the pattern of right, wrong, and omitted responses on all items administered to an individual student. Rather than merely counting rights and wrongs, the IRT procedures also considers characteristics of each of the test items, such as their difficulty, and the likelihood that they could be guessed correctly by low-ability individuals. IRT scores are less likely than simple number-right or formula scores to be distorted by correct guesses on difficult items if a studen'ts response vector also contains incorrect answers to easier questions. Another attribute of IRT that makes it useful for NELS:88 is the calibration of item parameters for all items administered to all students. This makes it possible to obtain scores on the same scale for students who took harder or easier forms of the test.

The previously-used plan of testing the longitudinal component of the field test sample in the same two subject areas year after year was abandoned in the 1991 field test for two reasons. First, in order to amass sufficient observations to evaluate free response -- formats, a far larger sample size would be required than would be provided by the longitudinal cohort. The second consideration was the necessity of administering several relatively time-consuming free response questions to each student as well as multiple-choice questions in the same content area for the purpose of comparing results of the two formats. The large amount of material needed in each subject area did not allow administration of two subjects per student as had been done before.

Five forms of the field tests were designed in order to try out as many free response questions as possible in the areas of mathematics, science, and reading comprehension. The subject areas, number of items, and timings are shown in Table 3.10.1. Several constraints needed to be satisfied. Since the five forms would be spiralled, and administered to different students in the same room at the same time, section timings needed to be consistent across all forms in order to avoid confusion. Larger sample sizes were required for calculating item statistics on the multiple choice items than were necessary (or desirable, given scoring costs) for the free response. Having two mathematics and two science forms, each with the same set of multiple choice items but different free response items facilitated trying out the largest possible number of free response items while maintaining the sample sizes desirable for the evaluation of the multiple choice items. Several features of this design were

**Table 3.10.1**

| | Section 1 25 minutes | Section 2 17 minutes | Section 3 4 12-minute items |
|---|---|---|---|
| Form I (Math A) | 9 old multiple choice items,+ 11 new items, mostly geometry | 12 new m.c. items, more advanced topics | 4 free response math items, separately timed |
| Form II (Math B) | same 20 items as Form I | same 12 items as Form I | 4 different free response math items |
| Form III (Science A) | 10 old m.c. items, plus 14 new, more difficult | 11 old m.c. hist/cit/geog items, plus 14 new, more difficult | 4 free response science items, separately timed |
| Form IV (Science B) | same 24 items as Form III | same 25 items as Form III | 4 different free response science items |
| Form V (Reading) | 2 old reading passages (14 items) plus one new passage (8 items) | 2 new, more difficult passages with 18 items | 4 free response reading not separetely timed |

less than optimal, but accommodations were made. For example, testing history/citizenship/geography (HCG) items in the middle of the science form and having only one reading from (and thus only minimal sample size for the reading multiple choice items) were considered preferable to the further complications that would result from increasing the number of field test forms.

In addition to the test items themselves, the booklets contained a short background questionnaire. Sex and race were asked in order to test for possible item bias. Course-taking history in each of the tested subject areas, including detailed breakdowns of mathematics and science courses taken, was obtained for the purpose of relating responses on the test items to curriculum exposure.

**The Multiple Choice Items:** The primary mandate of the NELS:88 cognitive test battery is that it provide individually-reliable scores suitable for longitudinal measurement of gain. This requirement dictates that multiple choice tests, building upon the structure established in the Base Year and First Follow-Up, must be constructed for grade 12. Since the 1987 and 1989 field tests had included twelfth grade students, grade-relevant item statistics were already available for a large pool of multiple choice items. As a result, less than half of the testing time in the 1991 field tests (42 of the 85 minutes) was devoted to multiple choice so as to leave as much time as possible to explore the free response options. The multiple choice questions selected for the 1991 field test served three purposes. First, some new and more difficult items were added to extend the difficulty scale at the upper level and to provide curriculum-relevant items for students taking advanced courses. Second, some previously-used "linking" items were retained in order to anchor the calibration of the new items. Third, the whole multiple choice component would be available for comparison with the free response results.

**The Free Response Items:** A total of 20 free response items were included: eight in mathematics, eight in science, and four in reading comprehension, with a sample size of about 400 students taking each item. The diversity of the items used, with respect to subject area, content, format, and difficulty, should provide a substantial information base that will contribute to the decisions to be made in designing the grade 12 test battery.

The mathematics questions were developed according to the principles outlined in The Development and Validation of a Set of Mathematical Problem-Solving Superitems (Romberg, 1982). In fact, several of the NELS:88 field test items were borrowed, with permission, from the items used in the Romberg study, with modifications guided by the results he obtained. Other items were adapted from the California Assessment Program (California State Department of Education, 1989). As was described above, an important requirement for useful test items is that their difficulty be appropriate for the students' level of ability. NELS:88 is not a minimum-competency test -- which would only need to provide a pass/fail decision at a given point -- but a test that aims to achieve accurate measurement across a continuum of ability levels. A major concern with the mathematics items is targeting the difficulty level to the student when there is such a vast diversity of course-taking history in this subject area by grade 12. The approach taken in Romberg's study, and applied here as well, was to build multi-step items that begin by testing a very basic skill or concept, and build to greater levels of sophistication. This approach allows useful information to be gained for both low-ability and high-ability students within the same problem. The very high resource cost of free response items (12 minutes per item in this field test, compared to one-half to one minute each for multiple choice items; plus the costs of hand-scoring thousands of tests) demands that maximum benefit be derived from each item used. Item content ranged from course-related mathematics concepts and computations (such as equilateral triangles and area/perimeter) to practical applications (including car braking distance and a train schedule), with some problems incorporating features of both within their multiple steps.

Several item types were used in the science free response tests. The first, called "figural response" requires the student to draw an arrow on the diagram shown to indicate the direction of movement, blood flow, etc. that would occur in the situation presented. This type represents a sort of hybrid of multiple choice and free response, with some of the advantages and disadvantages of each. Like a multiple-choice item, it is relatively quick to administer and quick to score (although for most applications, computer-scoring is not yet practical) and has a reasonably unambiguous correct answer. Like free response, it does not offer the student clues or the opportunity to guess one of a limited number of options. But figural response items offer, at best, only hints about the students' reasons for responding as they did. And the guessing issue is somewhat controversial. Although no explicit options are presented, in many cases there are obvious choices for the directions of arrows, so the format is not truly "open-ended." Other science free response questions used called for the student to draw a diagram, write an explanation, solve a problem, or modify and interpret a graph. Several of the items were based on subject matter suggested by members of the NELS:88 Technical Review Panel. Their content ranged from strictly course-related material to very general information that the student might have acquired in any of a number of different courses, or by exposure to science-related issues in the news media.

An objective of the reading comprehension items was to make them interesting enough that the students would be motivated to put their best effort into responding to what was, for them, a low-risk test. Questions were designed that would be relevant to the students' own interests and experience. For each item, the students read a short passage and then responded by writing an explanation, listing steps in the process described, or drawing a picture of the scene portrayed in the passage. Creating a free response task that is appropriate to the ability levels of all students in the sample (assuming some minimal level of reading and writing skills) is probably more feasible in reading comprehension than in it is in mathematics or science. Even if students were not able to do a good job of interpreting the information and producing a logical and well-written essay, they should all at least be able to understand the question and attempt a response.

After each of the four free response items in each test booklet, the students were given a short series of questions to assess their experience in answering the item: whether they had found it too hard or too easy; whether the time allotment of 12 minutes was too much or too little; whether the question was clear or confusing; and whether they gave the best answer they were capable of giving. Student responses to these questions will be used in selecting and modifying the format and content of the items for further use.

**Scoring the Free Response Items:** There are two types of scoring approaches typically used to evaluate free response questions: holistic and analytic. Holistic scoring assigns a single score that takes into account the overall impression or quality of the response. Analytic scoring rates each of a number of features separately, for example, using the correct equation, doing computations accurately, using the correct metric, and labeling variables. The analytic method was chosen to score the NELS:88 field test since it would give the maximum amount of information that could be used in evaluating the free response items for use in the grade 12 main study test. Analytic scoring also seemed particularly well suited to the well-defined, objective scoring criteria necessary for the measurement objectives of NELS:88.

The test development specialists who had written the field test items reviewed samples of student test booklets to identify distinguishable categories of responses for each step or feature of each question, and prepared detailed scoring guides. Groups of experts in each subject area (primarily high school and college teachers) were assembled, and scoring sessions were held in late March and early April. Working with one free response question at a time, the test developers explained the purpose of the item and scoring procedures to the readers, and gave examples of student responses that would typify different score categories. Scoring guides were modified before and during the scoring sessions in order to take advantage of the teachers' perspective on the material, to clarify distinctions between categories, and to incorporate new categories as unanticipated student responses were found. A ten percent sample of the test booklets was selected for a second reading by a different scorer for the purpose of evaluating scoring reliability.

After scoring was completed for each one of the free response questions, a brief discussion was held to elicit readers' evaluations of the success of the question and their comments on students' responses. Readers were asked for suggestions for improving the content and presentation of questions; they proposed modifications that would make the scoring guides clearer, more complete, and easier to apply; they commented on the value of each question in the context of what was important for a student to learn; and they evaluated the extent to which free response format was advantageous in being able to judge student performance.

For the purpose of data analysis, a single provisional scaled score for each free response item was built from the set of categorical analytic scores. Later sections describe the relationship of these scaled scores to student background and course-taking variables, to the multiple choice sections of the test, and to student evaluations of the test items. Tables 3.10.2 to 3.10.6 in Appendix C give brief descriptions of the content and purpose of each free response item and descriptions of the scale points. Appendix C also contains the complete text of the question, the scoring guides which define the categories for each analytic score, and the specifications for converting the analytic scores to a single scale for a sample of items.

### 3.10.3 Analysis of Cognitive Field Test Results: Multiple Choice Sections

**Sample Description:** Field test booklets were processed for 2,070 students. Sample sizes and breakdowns by sex, race, and course-taking history are shown in Table 3.10.7. Two different subsamples were selected for further analysis. Counts of omitted items were based only on students who appeared to have attempted the test, as defined by answering at least three items in each of the two multiple choice sections in the booklet. All other analyses (including item analysis of the multiple choice sections and analysis of free response questions) included only those students who had answered at least half of each multiple choice section in the subject area addressed by the analysis.

As can be seen by the numbers in the table, very few observations were lost due to students choosing to omit whole (or nearly whole) sections. The largest such loss was for the mathematics forms, in which about six percent of test takers failed to answer enough items to be included in the analysis. The high difficulty level of the questions in the second multiple choice section may have caused students who had not taken advanced courses in mathematics to become discouraged and not continue. Sample attrition rates were about the same for white and minority students, and for males and females.

Item analysis statistics for the multiple choice sections of the tests are presented in tables 3.10.8 through 3.10.11 in Appendix C. The first column of each table contains the r-biserial for the test items, which is an adjusted correlation coefficient relating an item to the total test score. High r-biserials indicate items that are measuring the same underlying factor as the test as a whole. Low coefficients suggest flawed items (such as two correct answers), items that are too easy or too difficult for virtually all of the test takers, or the possible presence of a different factor being measured. The items whose numbers are followed by an "X" in the tables were identified on the basis of low r-biserials, and were found either to have concepts or terminology that appeared to be unfamiliar to the test takers or to have more than one answer that tended to be chosen by the highest-scoring students. These questions will be removed from the item pool for the grade 12 tests unless their flaws are obvious and correctable. They have also been deleted from the field test scores for the remainder of the statistical analyses presented in this report.

Table 3.10.7

| | Mathematics | | Science/History | | Reading |
|---|---|---|---|---|---|
| | Form 1 | Form 2 | Form 3 | Form 4 | Form 5 |
| Total N | 414 | 450 | 412 | 371 | 423 |
| | | | | | |
| Male | 219 | 213 | 188 | 178 | 204 |
| Female | 187 | 229 | 212 | 188 | 212 |
| | | | | | |
| Black | 59 | 62 | 78 | 54 | 59 |
| White | 225 | 254 | 208 | 197 | 232 |
| Hispanic | 68 | 70 | 78 | 73 | 85 |
| Asian | 36 | 40 | 21 | 26 | 32 |
| American Indian | 4 | 2 | 1 | 1 | 1 |
| Other Race | 12 | 12 | 12 | 11 | 6 |
| | | | | | |
| Algebra (1+ year) | 88% | 87% | | | |
| Geometry (1+ year) | 73% | 77% | | | |
| Trigonometry (½+ year) | 50% | 49% | | | |
| | | | | | |
| Biology (1+ year) | | | 91% | 92% | |
| Chemistry (1+ year) | | | 57% | 57% | |
| Physics (1+ year) | | | 28% | 33% | |
| | | | | | |
| History (2+ years) | | | 91% | 91% | |
| History (3+ years) | | | 72% | 69% | |
| | | | | | |
| English (3+ years) | | | | | 94% |
| | | | | | |
| N for Omit Counts | 405 | 431 | 411 | 367 | 415 |
| N for Test Analyses | 395 | 416 | 409 (sci)<br>411 (hst) | 368 (sci)<br>366 (hst) | 411 |

Source: NELS:88 Second Follow-Up Field Test Data

The limited testing time available for multiple choice items in the 1991 field test meant that primarily new and more difficult questions were tested, not the full range of items that will be drawn on in constructing the 1992 tests. Items from the 1989 field test will be selected for the 1992 tests as well.

Neither of the field test samples was designed to be representative of the population, nor can they be assumed to be equivalent to each other in average ability. Proportion correct (P+) for the 1991 items is presented as well as the 1989 statistics for the subset of items that had been used in the earlier year. The difference in P+ for the same items tested in the two samples provides an adjustment factor to compensate for different ability levels of the groups. For example, for the 14 reading comprehension items that were used in both years, the 1991 sample scored, on average, eight percentage points higher than the 1989 group. Subtracting .08 from the P+ for each of the new items recalibrates the statistics so that they are roughly comparable to the scale of the rest of the item pool. The adjustment factors for the other three tests are somewhat smaller: .03 for mathematics, -.03 for history/citizenship/geography (HCG), and close to zero for science.

The item analysis tables also list the source and content of each item. All of the new reading comprehension and science items were written for the current field test. The new HCG items were selected from NAEP. New mathematics items were borrowed or adapted from NAEP and from SIMS (the Second International Mathematics Study). The content classification "Comparisons" for the new reading items refers to the fact that these items were based on pairs of reading passages, followed by sets of questions that ask the student to compare the two passages with respect to some substantive or stylistic feature.

The tests were not speeded, with over 93 percent attempting the last item in all subjects except mathematics, where 83 percent responded to the last item in section 1 and 91 percent finished section 2. This is an indication that the moderately difficult word problems that made up the 1991 mathematics test take longer, on average, than the mix of word problems and relatively fast quantitative comparisons used in earlier test forms. Overall, response rates were quite high, with an average of a little more than one omitted item in each subject area, slightly higher for the mathematics test and for Black students. See Tables 3.10.12 to 3.10.20 in Appendix C for breakdowns of "omit rates" by subject area, gender, race/ethnicity and course-taking background.

Variations in the size of the alphas primarily reflect discrepancies in the number of items in each test rather than differences in internal consistency. Simulations demonstrate that if all the tests were of equal length, comparable reliabilities would be attained.

The possibility of item bias was investigated by correlating item scores with dummy variables for gender and ethnicity, while holding total test score constant. On each test, about one to four items were found to have statistically significant, but small (in the range of .10 to .15), partial correlation coefficients for each of three contrasts: males vs. females, Blacks vs. Whites, and Hispanics vs. Whites. These items tended to be about evenly divided between those that favored each focal group and those that appeared to be biased against it. All identified items will be reviewed, and deleted from the item pool if the evidence of bias appears to be substantial.

Factor analyses were performed for each of the multiple choice tests. Cross products matrices were corrected for guessing, and tetrachoric correlations were employed. All tests proved to be strongly single-factor: the ratio of first to second roots was 4.8 for the mathematics test, 4.2 for science, 8.1 for reading comprehension and 3.5 for history/citizenship/ geography. When additional factors were extracted, factor loadings clustered by item difficulty rather than content for all subtests, strengthening the evidence that a dominant single factor was present.

Internal consistency reliabilities were computed for each of the four subject areas after flawed items had been deleted and are listed below in Table 3.10.21. Coefficient alphas were acceptably high, given the number of items and the skewed distribution of difficulties.

**Table 3.10.21**
**Reliabilities of Individual Tests**

| | # Items | Coef. $\alpha$ |
|---|---|---|
| Reading Comprehension | 38 | .86 |
| Mathematics | 28 | .78 |
| Science | 22 | .65 |
| History/Citizenship/Geog. | 24 | .74 |

Source: NELS:88 Second Follow-Up Field Test Data

Means and standard deviations of test scores are not presented here. Each subject area contains groups of items being tried out for a particular purpose, but none of them is designed to stand alone as a coherent and well-constructed test with a complete range of content and difficulty. However, since each student received multiple choice items in the same subject area as the four free response questions, and since there is evidence of similar factor structure, these scores will serve as a reference point for examination of the results obtained in the free response sections.

### 3.10.4 Data Analysis of Free Response Items

Tables 3.10.2 to 3.10.6 in Appendix C contain summaries of the content and purpose of the 20 free response items. Brief descriptions of the scaled scores are included, as well as the percentage of students receiving each score. Details of all of the individual analytic scores for each item, and the procedures for combining them to construct the scales, are too lengthy to be included here in full. Examples of these materials for three selected items are in Appendix C.

It should be noted that while we assume that the scaled scores indicate a monotonically increasing level of ability, there is no claim that the scale points represent equal-interval steps. For any items selected to be included in the 1992 survey, the appropriateness of the score scale (including a check on the monotonicity assumption, using multiple choice items as an anchor) will be investigated further and modifications made if necessary.

**Timing:** The free response items were "paced", that is, separately timed, at 12 minutes each. With the open-ended format, there is the potential for students to get bogged down in writing a much more complex response than was anticipated, and thus jeopardize their ability to finish the rest of the test. It then becomes impossible to tell whether unanswered items at the end of the test were too difficult or whether the student simply ran out of time. To avoid this problem, the field test students were told when the time was up for each question, and were instructed to move on to the next one. Tabulations of the student reaction questions showed that the 12 minutes allotted for each question was more than enough. Only about 10 percent of respondents said they could have used more time, while over half said they had more or much more time than was needed. These results were remarkably consistent over all 20 items. This is probably a good indication that the timing could be shortened to ten minutes per item. However, we still feel that pacing the items is important to ensure that each item has been presented to each student, whether or not the student chooses to respond.

**Nonresponse:** The tendency for students to omit items was substantially higher for most of the free response questions than for the multiple choice sections of the field test forms. Nonresponse rates for the multiple choice sections were about 4 to 5 percent overall, slightly higher for the mathematics sections (6 to 7 percent) and somewhat lower (2 to 3 percent) for reading comprehension. Five of the 20 free response questions had omit rates that were roughly comparable to the multiple choice sections in the same subject areas; the rest were considerably higher. Students were least likely to answer the science questions (omit rates of 11 to 59 percent) and most likely to respond to those in reading comprehension (5 to 12 percent nonresponse). Omit rates broken down by gender and ethnicity are presented in Tables 3.10.12 to 3.10.16 in Appendix C. Rates for students who had taken related mathematics and science courses are in Tables 3.10.17 to 3.10.20 in Appendix C. Each of these tables also contains summary statistics on the students' evaluation of the difficulty of the item, based only on those who omitted the question.

The mathematics free response items had been designed to minimize nonresponse by taking a multi-step approach within each item, that is, starting with a relatively easy question or concept and building to a more complex level. Thus students who could not complete the entire item might at least be able to provide enough information so that they could be placed on the score scale. This approach was partly, but not entirely, successful in reducing nonresponse. The items that were most closely course-related tended to fare the worst, even for students who had taken the relevant courses. A geometry question (Form 1, Q4) that was omitted by 34 percent of the students overall was not answered by 25 percent of the students who reported taking a year or more of geometry. The failure of students to respond, even those who had been exposed to related coursework, was true to an even greater extent in science, where most of the questions were not multi-step. A question concerning conservation of momentum (Form 3, Q3) was omitted by 23 percent of students who had taken a year of physics (43 percent of the total group omitted this question); an ecology question (Form 4, Q3) was not answered by 16 percent of students with coursework in biology (18 percent overall); and 36 percent of chemistry students did not respond at all to a question (Form 4, Q4) on balancing a chemical equation (59 percent overall). The best results, from the standpoint of nonresponse, were obtained with the reading comprehension items. Since these items drew on real-life situations, and since virtually all students were probably exposed to reading-and-writing exercises in their coursework (over 93 percent reported having taken three or more years of English in high school), these questions were by far the most accessible to the test takers.

**Gender and Race Bias:** An even more serious issue than the relatively high overall level of omitted items in the free response sections is the problem of differential impact by subgroups. While males and females tended to have roughly similar omit rates (with females generally slightly more likely to answer mathematics and reading comprehension questions, but somewhat less likely to respond in science), this was not the case for racial subgroups. Both Black and Hispanic students had substantially higher nonresponse rates than whites on virtually all of the free response items. For several items, minority nonresponse exceeded that for white students by a factor of two to three times, or by 30 or more percentage points for other questions. (By contrast, minority group omit rates for the multiple choice sections of the tests were within a percentage point or two of rates for the white students for all subject areas.) To some extent, higher nonresponse for Black and Hispanic students might be explained by differential curriculum exposure. But that would not account for differences in the reading comprehension form, where minority nonresponse rates (although much lower than for mathematics and science) were still higher than for white students even though virtually all had taken English courses throughout high school.

For the students who did answer the free response questions (or indicated that they were unable to), differential impact by subgroups remains a problem. When free response scale scores are correlated with a dummy variable for minority group membership, Black and Hispanic students perform significantly worse than white students on all items. This is consistent with results elsewhere, and does not in itself constitute item bias. A much more serious finding is that the correlations remain negative for 13 of the 16 mathematics and science items, and one of the four reading comprehension items, even when total score on the multiple choice test is controlled for. Although the field test sample is not representative of the grade 12 population, the consistency of these results suggests that careful study of possible factors accounting for the apparent item bias is in order.

A consistent pattern of male/female item bias was not found. After controlling for total test score, none of the eight mathematics free response items favored either gender significantly. The remaining 12 science and reading comprehension items were nearly evenly split between those that favored males, those that favored females, and items on which neither gender had an advantage. Correlations of scaled scores with ethnicity and gender are shown in Table 3.10.21.

**Imputed Scores:** The Item Response Theory model used to score the NELS:88 multiple choice tests in the past satisfactorily compensates for missing data, particularly if the tests are strongly single-factor and the nonresponse rates are low. There are partial-credit models designed to be used for continuous score scales, but identifying the appropriate model for NELS:88 (if one exists) and modifying the rest of the test structure if necessary to meet its requirements would not be a simple and straightforward process. An ad hoc method for imputing scores for some of the missing data was used for the field test data analysis and is described below.

In a motivated test (such as a college entrance exam or a high school proficiency test, where it is in the test taker's interest to get a high score), it may be assumed that an unanswered item is equivalent to an incorrect response. But in our field test of free response items, there is some evidence that lack of motivation rather than lack of ability was responsible for some of the nonresponse. Many test-takers left an item blank even though they either reported having taken the courses that would have prepared them to attempt it, or indicated in the subsequent student-reaction question that they had found the item to be easy or about the right difficulty. Others made irrelevant or inappropriate remarks instead of answering the question. In a classroom situation, this type of nonresponse might appropriately be assigned a failing grade. But for NELS:88 the issue is accurate measurement, not pass or fail. The assumption that missing data is always due to inability to answer would produce misleading results in analyzing the effects of educational processes in producing achievement outcomes.

Nevertheless, for over half the unanswered items, the student indicated having found the item "hard" or "too hard." For these cases we assumed that the student was not able to answer the question and we provisionally imputed the lowest point on the score scale for the purpose of computing correlations, reliabilities, and factor analyses. As a check on the reasonableness of the imputing procedure, we referred to the multiple choice test score and the self-report of coursework to see if there was evidence that the students we assigned to the lowest scale point might really have been able to perform at a higher level had they attempted the item. In general, the data supported the assumption that most of the students who omitted an item and indicated that it was too hard probably would have scored poorly. Table 3.10.23 in Appendix C contains counts of imputed scores for each free response item, along with the numbers of these students who either scored in the top half of the distribution in the multiple choice test, or took one or more courses in a related subject, or both.

However, there are several considerations that argue for further investigation. For some of the free response questions, there were substantial numbers of high-multiple-choice-score students who omitted the item and said it was too difficult. The coursework questions used here, while probably accurate for most students, only report how many courses each student took in related subjects, not whether he or she successfully mastered the material. In addition, the current imputing method does not address the problem of estimation for students who did not answer the difficulty question, or claim to have found the item easy or appropriate difficulty but omitted it anyway. If score imputing is to be used for estimating missing free response data in 1992, additional evidence must be obtained for these less obvious cases.

**Reliability:** Cronbach's alpha is the most commonly used statistic for assessing internal-consistency reliability. It uses individual item variances and total score variance to calculate the proportion of "true variance" in the test. Unanswered items can distort the coefficient alpha by artificially raising the total score variance. There are two possible ways of compensating for this problem. In a motivated (high-risk) test, omitted items may be treated as if they were answered incorrectly; or if the total amount of missing data is very small, only complete-data cases may be used to compute the alpha. We did all computations for students who had answered (or received imputed scores for) at least two free response questions, and again for those with complete data on all four. The two alternative methods produced very similar results.

Another perspective on reliability is that students should receive comparable measurements on different forms of a test. Split half reliability simulates two different test forms by scoring each half of the test separately and then correlating the results. An adjustment compensates for the fact that each score is based on a half-length test.

The NELS:88 test was neither a high risk test, nor was there a negligible amount of missing data in the free response items. Three approaches were taken to estimating the increased reliability that could be achieved by adding free response items to the multiple choice test: imputing scores for some of the missing data, as described earlier; computing alphas based on both complete and incomplete free response sections to estimate lower and upper bounds for the reliability; and computing split-half reliabilities as well as alphas.

For the mathematics and science forms, the addition of free response items increased the test reliability; for reading comprehension, reliability was lowered slightly by the addition of free response items. (Part of this finding is related to the fact that the multiple choice reading test with 38 items was already the longest, the most reliable, and the most strongly single-factor of the four subject areas; the science test, in which only 22 items were analyzed, had the lowest reliability and thus the most to gain from the additional measures.) Alpha coefficients for the multiple choice sections alone were .78 for mathematics, .65 for science, and .86 for reading comprehension. A single free response item added one to three percentage points to the alpha for mathematics (except for one item, which lowered the reliability slightly); three to five percentage points for science (again, except for one item); and lowered the reading comprehension reliability by up to three points. Adding all four free response items provided an increment over multiple choice alone of .04 for mathematics and .10 for the short science test, while it lowered the alpha by .03 for reading. The improvement in split-half reliability showed a similar pattern: a seven point gain for mathematics (from .77 to .84 when all four free response questions were added), a ten point gain for science (.67 to .77) and a one point loss for reading (.86 to .85).

For comparison purposes, the Spearman-Brown Prophecy Formula was employed to estimate the increment in alpha that could be achieved by adding more multiple choice items to the test instead of free response. In both mathematics and science, five more multiple choice items would increase the reliability about as much as the best single free response item. The comparable number of additional multiple choice items required to match the increment contributed by all *four* free response items was about eight for mathematics and 14 for the science test. In each case, the additional multiple choice items would take less additional testing time than the corresponding free response and would not entail the costs of hand-scoring that are necessary for free response format.

Of course, free response items can serve other purposes than merely adding to the reliability of a test in a cost effective manner. The score scales constructed for this analysis were deliberately designed to try to measure the same construct as the subject area tests. For example, we attempted to score a mathematics item to reflect only the student's mathematics knowledge and not other factors such as writing ability, neatness, or skill in drawing. But this objective was arbitrary. Which individual features of a response were chosen for scoring, and the way in which the features were combined to build a scale, could have been defined in many other ways. A rule of thumb in test construction at ETS is that multiple choice items serve the purpose of content coverage, and free response is the appropriate format for a "work sample." The types of skills the work sample is intended to measure determine the criteria appropriate for designing the scoring scheme.

The free response items have an additional source of potential error variance that is not present in multiple choice tests: the imperfectly-objective judgments of the readers. Reader reliability (which is the extent to which different readers would give the same or different scores to a particular answer) imposes an upper limit on the reliability that can be achieved by the item. A ten percent sample of the free response items was selected to be scored by a second reader in order to assess this component of total reliability. For most but not all of the items reader agreement exceeded 80 percent when the scaled score was constructed from the individually-scored features. This can probably be improved somewhat by identifying the specific analytic scores that had the lowest reader agreement and clarifying scoring instructions for these features. Tables 3.10.24 to 3.10.28 in Appendix C present statistics on reader agreement on each of the categorical analytic scores, as well as crosstabs of Reader 1/Reader 2 scaled scores when the individual features are combined to build continuous scales. The "na" category in the tables includes omitted items and inconsistent responses that could not be scaled.

**Factor Structure:** When free response items were added to the factor analysis described in the section on multiple choice results, neither a guessing correction nor tetrachoric correlations could be employed since these items were scored on a continuous scale rather than right/wrong. For comparison purposes, the multiple-choice-only factor analyses were repeated without the guessing and tetrachoric adjustments, and the first-to-second root ratios before and after the addition of the four free response items were compared. For mathematics and science, the free response items increased the ratios by about 22 - 26 percent, for reading comprehension only 2 percent. Virtually all of the free response items had their highest loadings on the first factor. (Zero-order correlations of the free response items with multiple choice test total scores were also uniformly high, averaging .52 for mathematics, .44 for science, and .39 for reading comprehension; again indicating a very strong relationship between the constructs being measured by the two item formats.)

Had the scoring scales been designed differently, for example, giving weight to features such as writing style, other factors might have emerged. Just as free response format provides test takers the opportunity to respond in many different ways, it licenses multiple and sometimes arbitrary evaluation criteria.

**Student and Reader Reactions to the Free Response Questions:** Test takers were asked four questions after each free response item: how hard they had found the question, whether they had enough time, whether the question was clear or confusing, and whether they gave the best answer they could. The first two of these questions provided valuable information relevant to timing and score imputing, as discussed earlier. The last two were less successful. Students tended to answer that the question was clear if it was of appropriate difficulty, and to say it was confusing if it was too hard for them. Thus they were essentially assessing their own ability to answer, rather than evaluating the clarity of the item as intended. Similarly, the last question was meant to assess the extent to which answers were poor due to lack of motivation rather than lack of ability. On most questions, about 70 to 80 percent of students said yes, they gave the best answers they could. But only a few of the "no" answers addressed motivation (e.g., "No, because this is a waste of time."). Most of the others said something like, "No, because I haven't had a course in this subject," again assessing their own skills. This type of response was much more frequent for the technical science questions. In fact, these students had given the best answer they could -- they simply did not have the background to do any better. Fortunately, the readers who scored the questions were able to provide some of the feedback these questions had been designed to elicit, such as identifying aspects of some of the questions in which the students did not clearly understand what was expected of them.

Brief discussions were held with the readers after each question was scored, during which they offered suggestions for improving the questions and scoring guides. They were also asked whether each question was worthwhile in testing valuable concepts and whether free response format was advantageous relative to multiple choice. For most questions, the readers felt that the content was worth testing. The most frequently criticized items were those that tested material that would only have been learned in advanced courses that most of the students had not taken (for example, the dice-probability problem and the chemical equation). The readers questioned the value of administering these items to thousands of students for whom the items would be totally irrelevant simply because they had not been exposed to the required course material.

Free response format was thought to be advantageous for a majority of the questions, particularly those that involved writing a short essay or an explanation. However, readers believed that some of the questions, or parts of questions, could be just as successful in multiple choice format, particularly when the incorrect answers tended to cluster in a small number of patterns that could be presented as multiple choice options. This was true of the first parts of the dice problem, the conservation of momentum question, the positions of earth and sun in summer and winter, and all of the figural response items. Readers suggested the possibility of using both formats within an item: presenting the short-answer parts as multiple choice, and only using free response for the explanation step.

**Scoring Costs:** Costs of administering free response items in the 1992 survey are difficult to estimate precisely without knowing certain parameters such as the number of different items to be used overall, the particular items to be selected, and the number per student. Scoring times for the twenty field test items ranged from an average of about three-quarters of a minute per item for some items, to just over a minute and a half. The cost of scoring 8,000 free response items in the field test (four items each for 2,000 students) was roughly $35,000. This figure includes the costs of readers, test development specialists who developed scoring guides and conducted the scoring sessions, and data entry. It does not include development of the test items themselves. The cost per item might be somewhat reduced in the main study by printing a scannable scoring grid in the test booklet and by the economies of scale that could be realized by scoring large samples of a few items, rather than the small samples of many different items as in the field test. On the other hand, field test readers were recruited from within a local area, so travel and subsistence costs were minimal. Readers for larger scale scoring sessions would need to come from a wider geographical area, thus adding to the cost per item.

### 3.10.5 Recommendations for the 2FU Main Study Cognitive Test Battery

The following recommendations are driven by three equally important considerations: psychometric issues, cost effectiveness, and usefulness of the measures to users of the NELS:88 database.

We believe that the primary responsibility of the NELS:88 test battery is to provide reliable measures of achievement appropriate for measuring status at each point in time as well as gain over the four year time interval. Ideally, we would like to incorporate some experimental measures as well, but limited resources may dictate that NELS:88 cannot be all things to all people. It would be a disservice to our constituency -- the future users of the database -- to compromise the usefulness of the test battery in order to attain secondary objectives.

**Preliminary Specifications for the Grade 12 Cognitive Tests:** Multiple choice tests that follow the structure used in grades 8 and 10 would be best suited for measurement of change over time. Details to be resolved when the grade 10 test data have been analyzed include item selection and test length. Evaluation of the success of the multi-level reading and mathematics tests will determine whether this design will be retained in grade 12. The science test will probably need to be either multi-level or longer than in the past, in order to address the diversity of curriculum exposure in the student population.

**Mathematics:** Current plans include the possibility of putting NELS:88 grade 12 mathematics scores on the same scale as NAEP. The validity of this procedure depends on the comparability of the two 12th grade samples and of the test specifications. While the NAEP and NELS:88 12th grade samples are not strictly comparable, analytic procedures can be used to provide reasonably matched samples. Similarly, the present NELS:88 item pool can provide a reasonable match to the 1992 NAEP specifications. Table 3.10.29 presents the process and content specifications for NELS:88 grades 8 and 10, as well as the 1992 NAEP 12th grade mathematics specifications. The latter will serve as an approximate target for NELS:88 12th grade mathematics items, averaged over all of the NELS:88 multi-level forms. For example, we would anticipate that the number of items involving simple arithmetic operations would be less than 25 percent (the percent in NAEP).

**Reading:** In the case of reading, there is no present plan to "crosswalk" between NAEP and NELS:88. However, it would seem worthwhile to use NAEP 12th grade specifications as a preliminary guide. Table 3.10.30 presents the NELS:88 process (mode) and content (type) breakdown for NELS:88 grades 8 and 10. (The 1990 NAEP reading classification uses the term mode where we would use process, and type where we use content.) The 1990 NAEP test specifications are also shown in the columns on the right. The 1992 NAEP reading specifications have been so drastically changed that they cannot be considered as a reasonable target specification for NELS:88. It would be very difficult for NELS:88 to duplicate the NAEP 1990 12th grade specifications exactly since the number of items in the NELS:88 reading pool is quite limited. In addition, the NAEP reading specifications contain a "document" classification not present in NELS:88. Replicating the NAEP specifications would require adding "document" items to each of the NELS:88 forms. However, given the longitudinal nature of NELS:88, such a change could alter the factor structure and thus cause problems in the vertical scaling. We suggest that the 10th grade reading specifications be retained, with an additional, more difficult reading passage added to the higher level form.

**Science:** As in the case of reading, the 1990 NAEP Science test specifications provide some guidelines for the NELS:88 grade 12 test. Table 3.10.31 presents the NELS:88 8th and 10th grade specifications as well as those of the 12th grade NAEP item pool. As indicated above, we suggest that the science test either be made multi-level in grade 12, or increased in length from 25 to 30 items to increase reliability. If we decide to stay with one form, we might wish to add two or three physical science items and an additional single item in scientific method and earth science. If two forms are used, the 25-item length would be retained. The lower level form would have three fewer items in chemistry and three more in life science and/or earth science. The higher level form would have three more items in physical science and three fewer in life and/or earth science. Lengthening the form might be preferable to having two forms in the 12th grade for two reasons. First, there is the possibility that the unidimensionality of the science test may be threatened by the addition of higher-level curriculum-related items. The longer single form would allow us to address this question. Secondly, the multi-level design in science increases the number of different test forms needed (already six forms for the combinations of two reading and three mathematics levels), and thus the potential for logistical problems in labelling and distribution.

## Table 3.10.29

### MATHEMATICS FRAMEWORK FOR NELS:88 ASSESSMENT

(40 Items)

| | CONTENT | | | | |
|---|---|---|---|---|---|
| | Grade 8 | Grade 10 | | | NAEP Grade 12 |
| | | Low | Medium | High | |
| Arithmetic | 47% | 54% | 45% | 30% | 25% |
| Algebra | 28% | 28% | 29% | 30% | 25% |
| Geometry/ Measurement | 10% | 10% | 13% | 15% | 35% |
| Data Interpretation/ Probability | 10% | 5% | 8% | 13% | 15% |
| Advanced Topics | 5% | 3% | 5% | 12% | - |

### MATHEMATICS FRAMEWORK FOR NELS:88 ASSESSMENT

(40 Items)

| | PROCESS | | | | |
|---|---|---|---|---|---|
| | Grade 8 | Grade 10 | | | NAEP Grade 12 |
| | | Low | Medium | High | |
| Procedural Skills/ Knowledge | 42% | 47% | 35% | 33% | 30% |
| Understanding Comprehension (Conceptual Understanding) | 48% | 43% | 45% | 45% | 40% |
| Problem Solving | 10% | 10% | 20% | 22% | 30% |

**Table 3.10.30**

## READING FRAMEWORK FOR NELS:88 ASSESSMENT

(21 Items)

| | CONTENT | | | |
|---|---|---|---|---|
| | Grade 8 | Grade 10 | | Approx. NAEP Grade 12 |
| | | Low | High | |
| Literary | 38% | 38% | 43% | 10% |
| Science | 14% | 14% | 14% | 25% |
| Poetry | 29% | 29% | 24% | 10% |
| Biography | 19% | 19% | 19% | 30% |
| Social Studies | 0 | 0 | 0 | 25% |

## READING FRAMEWORK FOR NELS:88 ASSESSMENT

(21 Items)

| | PROCESS | | | |
|---|---|---|---|---|
| | Grade 8 | Grade 10 | | NAEP Grade 12 |
| | | Low | High | |
| Reproduction of Detail | 19% | 19% | 19% | 35% |
| Comprehension | 14% | 19% | 29% | 35% |
| Inference/ Evaluation | 67% | 62% | 52% | 30% |

**Table 3.10.31**

## SCIENCE FRAMEWORK FOR THE NELS:88 ASSESSMENT

(25 Items)

| | CONTENT | | |
|---|---|---|---|
| | Grade 8 | Grade 10 | NAEP Grade 12 |
| Life Science | 32% | 28% | 32% |
| Earth Science | 32% | 20% | 22% |
| Physical Science | – | 16% | 17% |
| Chemistry | 28% | 28% | 17% |
| Scientific Method (Nature of Science) | 8% | 8% | 12% |

## SCIENCE FRAMEWORK FOR THE NELS:88 ASSESSMENT

(25 Items)

| | PROCESS | | |
|---|---|---|---|
| | Grade 8 | Grade 10 | NAEP Grade 12 |
| Declarative Knowledge (Knowing Science) | 40% | 32% | 40% |
| Understanding Comprehension (Uses) | 28% | 28% | 20% |
| Problem Solving (Integrating) | 32% | 40% | 40% |

## Table 3.10.32

### HISTORY/CITIZENSHIP/GEOGRAPHY FRAMEWORK FOR THE NELS:88 ASSESSMENT

(30 Items)

| | CONTENT SPECIFICATIONS | | |
|---|---|---|---|
| | Grade 8 | Grade 10 | Grade 12 |
| American History | 47% | 63% | 45% |
| Citizenship/ Government | 43% | 27% | 45% |
| Geography | 10% | 10% | 10% |

**History/Citizenship/Geography:** In the absence of comparable specifications from NAEP, we suggest an approximately equal balance of history and citizenship-government items in the 12th grade year, with geography kept constant at about 10 percent of the test items. However, results from the 10th grade main study testing, when they become available, may show that the more difficult items tend to be history questions. If that is the case, then it might be appropriate to have proportionatly more history than citizenship items in grade 12 to minimize ceiling effects. Table 3.10.32 shows content distributions for the earlier years and the proposed proportions for 12th grade. Items were not classified by process for this test.

**Deleting Subtests:** We believe that deleting any of the four subtests from the 1992 survey to make time for free response items would be a serious mistake. We come to this conclusion not only from our perspective as builders of the NELS:88 database, but from our experience as users of the NLS-72 and High School and Beyond data. We have participated in a large number of studies based on the earlier surveys, and in most of them cognitive test scores have played a major role. Reliable measurement of achievement in grade 12 serves two essential purposes in NELS:88. First, it permits measurement of gain over the full period of the high school years, which was one of the reasons for beginning the NELS:88 survey in the eighth grade. Second, the senior year scores will serve as the ability measures for studies of higher education decisions that will be undertaken based on the post-high school follow-ups.

The reading and history/citizenship/geography tests have been suggested as candidates for shortening or deleting from the grade 12 test battery. It has been claimed that students gain little in reading comprehension skills during the last two years of high school, so the grade 10 test would serve as an adequate measure of achievement. The HS&B sophomore cohort test results contradict this contention. Overall, HS&B students gained about a quarter of a standard deviation in reading from grade 10 to grade 12, about the same as the mathematics gain. Moreover, the gain in reading achievement was not a constant for different ethnic groups, SES breakdowns, or school dropouts, exactly the sort of issues that studies of the NELS:88 data will want to investigate. In addition, the NELS:88's student sample will expand in 1992 to embrace "freshened" students and previously ineligible students whose eligibility status has changed, for whom no reading data would be available should the reading test be dropped.

We cannot address the value of retaining the history/citizenship/ geography test on the basis of its usefulness in prior surveys, since only the HS&B sophomore cohort was tested in this subject area. The HS&B civics test, which contained only ten items, was too short to achieve an adequate level of reliability, and we are not aware of it having been widely used. (This experience suggests that care must be taken to maintain a sufficient number of items for each of the subtests in the NELS:88 battery.) The investment so far in giving the history/citizenship/geography test in the eighth and tenth grades, and the fact that most students do continue taking coursework in these subjects in the later high school years, would argue for retaining this test in the absence of compelling reasons to the contrary.

Another issue to keep in mind is that the 1992 "freshened sample" will contain students who have not been tested in the earlier years. We will, of course, not be able to measure gain for these students. However, the grade 12 tests will serve as the only achievement measures available for this group.

**Combining Multiple Choice and Free Response:** If the reliability and factor structure of the tests were the only considerations, incorporating free response items as a part of the multiple choice test sections might not be an inappropriate thing to do. The items appear to measure the same constructs as the multiple choice tests (or at least constructs that are highly correlated) and they generally enhance the reliability of the tests more than would an equal number of additional multiple choice items. However, the increased reliability to be achieved by adding free response items is less than could be gained from equal time spent on multiple choice items, to say nothing of the additional scoring costs incurred. The much higher nonresponse rates for Black and Hispanic students relative to whites would also argue against combining multiple choice and free response items at this time. In addition, results from NAEP administrations suggest that inserting free response items in multiple choice blocks leads to undesirable context effects. That is, multiple choice items that follow a block of free response items tend to have higher nonresponse rates than those preceding the free response.

If time and money are available in the 1992 survey, including free response items that are scored and reported separately from the multiple choice tests would, in our opinion, represent a better use of resources. The greatest value to be gained from free response items in NELS:88 is the diagnostic information that could be derived from analysis of detailed aspects of students' responses, as measured by multiple analytic scores. Burying all of the separately-scored features of a student's response by reducing them to a single scaled score, or worse yet, a dichotomous right/wrong score, is a waste of the student effort and the scoring costs that went into generating the data. The resulting dichotomous scores, while easier to analyze from a psychometric point of view, can be more efficiently measured in other ways. The free response items could also serve as a research database for exploring patterns of responses, and of nonresponse, for different population subgroups. We suggest including questions on item difficulty, as was done in the field test, to provide a basis for experimentation with methods of compensating for nonresponse.

**Spiralling:** The advisability of spiralling free response items in the 1992 survey depends on the use to which they will be put. If the objective of including the items is to create a database for the study of the items themselves, spiralling items would provide the possibility of including more diversity of item types and content. It is still important that sample sizes for each item remain high enough to support subgroup analyses. However, if the objective is using free response items for measuring students, spiralling may be risky. The items used in the field test are not of equal difficulty, and are not designed to be parallel measures. Even if equated scores can be produced by using IRT methods with the multiple choice section as an anchor, a substantial amount of equating error may remain due to the imperfect correlations among the free response items.

**NELS:88 2FU Field Test Database:** If resources are not available for including free response items in 1992 (or perhaps even if they are), we would like to recommend that the NELS:88 2FU Field Test database be made available to researchers interested in exploring issues in free response testing. The analysis of the test results undertaken for this Field Test Report has focused primarily on issues relevant to the design of the NELS:88 test battery. Other purposes could be served by additional analysis of the data -- for example: studies of individual features of student responses; alternative methods of constructing score scales; in-depth subgroup analyses; and relating responses to subsets of multiple choice items or to background characteristics. An edited and documented database now exists for the 2,000 students who took the cognitive tests. It contains the multiple choice as well as the free response items, some background information (sex, race, and course-taking history), and student-reaction questions linked to each free response item. It has sufficient minority group members (54 to 79 Black students and 68 to 85 Hispanic students taking each of the five forms) to identify clear patterns of responses. In addition, each of these test takers participated in the field test of the student questionnaire, which would provide additional background and context information. The questionnaire data for the 996 respondents that were used for the Student Questionnaire Field Test analysis could easily be merged with the cognitive test data. Data for the other half of the sample which have not at present been data-entered, could be added at minimal cost.

A word of caution is in order if this database is released to researchers. The field test sample is not a systematic sample and does not claim to be representative of the grade 12 population. Researchers would have to be effectively discouraged from drawing inferences that would be inappropriate for this reason. However, many or most in-depth studies of comparable topics are not based on probability samples, so this limitation is far from being a fatal flaw. If the entire database can be made available to interested researchers, the money invested in the field test will have been well spent, independent of the uses made of the items in NELS:88 twelfth grade year.

## Chapter 4: Dropout Survey Procedures

This chapter describes and evaluates the data collection and preparation procedures for the dropout component of the First Follow-Up Field Test. Group survey administrations -- conducted by NORC field staff -- were held at central sites in the five field test states from February through March of 1991. The success of survey procedures is assessed on the basis of field staff reports and project staff evaluations. Specific recommendations are made for changes in these procedures, focusing on refinement of the dropout definition and its operationalization and methods to increase dropout participation in the group survey sessions.

Before we begin the discussion of 2FU data collection and preparation procedures, it may be useful to consider the prior definitions of dropouts, including the sampling and questionnaire assignment issues, from the experience of both the HS&B and NELS:88 First Follow-Up studies.

### An Historical Perspective: Defining Dropouts in NELS:88 1FU and HS&B 1FU

Since the study of dropouts is a critical feature of NELS:88, it may be useful to summarize the methodological approach to this component taken in the NELS:88 First Follow-Up as well as the methods followed by High School and Beyond. Both studies form clear precedents from which the Second Follow-Up must learn. The HS&B sophomore cohort First Follow-Up (1982) and the NELS:88 Second Follow-Up (1992) both were designed to identify and study dropouts two years after the spring term of the sophomore year. NELS:88 First Follow-Up located and surveyed early dropouts -- those who left school between the eighth grade and the spring term of sophomore year -- these early dropouts will be included in the NELS:88 2FU. In addition to learning from the experience of these prior studies, the Second Follow-Up must seek some level of procedural and definitional consistency with them in order to achieve longitudinal and cross-cohort comparability.

**In School versus Out of School**: There are two levels of definition at which the basic distinction between in school and out of school must be applied: a classificatory level (a sample member is to be classified as a dropout or former dropout [stopout] or a student) and a data collection level (who should complete the dropout questionnaire?; who should complete the student questionnaire?). The classificatory level carries with it a sampling implication. Dropouts are retained with certainty in NELS:88; students are subsampled.

**Dropouts, stopouts, and chronic absentees**: Apart from regular students, there are three primary groups of interest for NELS:88 1FU: **cohort dropouts**--former students who are out of school in the spring term of 1990; **stopouts** (temporary dropouts, who are back in school in the spring term of 1990, but had a dropout episode between spring term 1988 and spring term 1990); and **chronic truants** (students who do not meet the conditions of the formal dropout definition, but have an exiguous physical presence in the classroom). Each of the three populations of interest: chronic truants, stopouts, and dropouts -- may be considered in turn.

**Chronic absentees**: Because NELS:88 1FU pursued substantial numbers of absent on Survey Day/absent on Make-Up Day sample members, item 13 in the 1990 student questionnaire may be of some value in identifying chronic absentees. (This item reads: in the first half of the current school year, about how many days were you absent from school for any reason? Response options range from "None" to "21 or more" -- 21 or more constitutes the number of absences required to be classified as a dropout or stopout).

**Stopouts:** Two issues must be confronted here: who is a stopout?; what questionnaire (in school [student] or out of school [dropout]) should the stopout complete?

Dropouts were identified in NELS:88 1FU through checks on enrollment status at three distinct stages subsequent to the spring term of 1988:

Phase 1: Tracing; spring term 1989 (eighth grade cohort members traced and enrollment status ascertained).

Phase 2: Autumn school contacting; fall 1989 (verifying sample members'school enrollment, freshening the sample).

Phase 3: Data collection; spring term 1990 (reverification of school enrollment status).

Hence, a sample member could have been identified as a dropout at phase one or phase two, but identified as back in school at a later phase. (A similar three phase scheme for ascertaining school enrollment status will be employed in the Second Follow-Up.)

In contrast, stopouts were identified in HS&B through an item on the student questionnaire (Q. 17) that asked if there were gaps in school attendance of a given duration (such as four consecutive weeks of unexcused absence).

The weakness of the NELS:88 1FU approach is that some stopout episodes of a brief duration may be missed. The weakness of the HS&B approach (apart from the fact that 1FU Q. 17 made no attempt to place the stopout event in time) is that it depends on a student report, which in turn assumes that the student is a respondent, and truthfully completes the item.

The NELS:88 Second Follow-Up will capture both within-wave stopout spells (say, a student who drops out in autumn of 1990 but returns to school in autumn of 1991) and cross-wave stopout episodes (a 1FU dropout who is enrolled in school in the 2FU). It should be considered whether the risk of missing some brief stopout spells justifies adding a question akin to that asked in HS&B 1FU, anchored to the period spring term 1990 to spring term 1992.

The second issue concerning stopouts is what instrument they should complete. In HS&B, as in-school students, stopouts received the student questionnaire. In NELS:88 1FU, this model was followed with one modification -- stopouts who had just returned to school were administered the dropout questionnaire. While there is a real loss in not obtaining from stopouts some of the data on the dropout experience that is contained in the out-of-school questionnaire, it is difficult and fundamentally inappropriate to administer to a student a questionnaire designed for sample members who are currently out of school. For the NELS:88 2FU main study, it is proposed to maintain this policy of administering the student questionnaire to stopouts who have been back in school for two or more weeks, and the dropout questionnaire to those who have been back in school for a lesser time period. At the same time, from a sampling perspective, it is proposed that stopouts be treated as a special category of dropout and be retained with certainty rather than being subject to subsampling. "Once a dropout always a dropout" would be the sampling principle applied in the Second Follow-Up.

**Dropouts:** In NELS:88 1FU, it was necessary to operationalize the concept of a dropout at two basic levels. First, in terms of an individual's status, who is a dropout and who is not? Second, in terms of instrument completion, who should be given the dropout questionnaire? Thus, in the NELS:88 1FU, all sample members who were out of school in the spring term of 1990 were to be administered the dropout questionnaire, but dropouts identified at earlier stages who had been back in school for some time (stopouts) were asked to complete the student questionnaire.

In the NELS:88 First Follow-Up, the following dropout definition was used:

1. an individual who, according to the school (if the sample member could not be located), or according to the school and home, is not attending school (or, more precisely, has not been in school for four consecutive weeks or more and is not absent due to accident or illness)

2. a student who has been in school less than two weeks after a period in which he or she was classified as a dropout (this individual would be classified as a stopout but administered the dropout, rather than the student, questionnaire; all other in-school stopouts would be administered the student questionnaire)

Note that this definition requires that dropout status normally be double-confirmed: both the school and the household must agree in their reports that the sample member's school attendance behavior conforms to the NELS:88 definition of a dropout (for further detail on implementation of this method of double confirmation, including decision rules for adjudication of conflicting reports, see Ingels and Scott, 1991). It is proposed to maintain in NELS:88 2FU these same definitions of a dropout both for classification purposes and for administration of the out of school questionnaire. It is also proposed to retain the basic methodology of double confirmation -- school and household -- of dropout status. It should be noted, however, that apart from what the school and household think the sample member's enrollment status to be, the subjective consciousness of the sample member is also a factor that must be considered, especially in the assignment of forms. The question of who should complete the student questionnaire, versus who should complete the out of school questionnaire, was a major issue in the NELS:88 1FU, as in HS&B before it.

While a basic definition of dropping out can readily be constructed, it is not easy to assign univocal meaning to the phrase "in school." This is particularly the case owing to the diversity and prevalence of alternative programs. Continuation schools and programs providing for students whose education in regular high schools has been (or -- an important distinction -- might be) interrupted render the concept of school enrollment status somewhat ambiguous. It is important to ask in such cases both what conceptual distinctions make sense, and how individuals in various categories of alternative programs -- dropout prevention, dropout re-entry, maternity centers, adult night classes for GED conducted at a community college -- actually think of themselves, that is, do they view themselves as enrolled in high school?

NELS:88 1FU considered students enrolled in alternative programs as students rather than dropouts (both for sampling and questionnaire administration), regardless of the nature of the alternative program. In the NELS:88 First Follow-Up field test in the spring term of 1989, it was found that when students in alternative programs were asked to complete the dropout questionnaire, oftentimes they found it difficult to answer some items because these questions implied that they had left or were not in school. Many of these students reported that they were not dropouts and still were students, even though in several of these instances they could not have been referred to their alternative program unless they had dropped out of school. (The degree of self-definition as a student felt by such sample members may possibly vary with the nature of the alternative program or with the way that school is defined by data collectors, and may be more intensely felt by early dropouts, especially those below legal school-leaving age, than by older dropouts -- this matter is worthy of further investigation). Our conclusion was that there may be some reluctance to identify oneself as a dropout when one is a participant in an alternative program, and that the student questionnaire -- if one is limited to but two questionnaires -- may be the more appropriate survey instrument for alternative program participants to complete. It should be noted that participants in alternative schools and programs were regarded as students for sampling purposes as well -- these students were not retained with certainty. Only dropouts (and institutionalized sample members) were retained with certainty in the NELS:88 1FU (although Asians, Hispanics, and American Indians were assigned higher probabilities of retention than other sample members).

Certainly the population of students who are in various degrees of disengagement from school is highly differentiated. There are students who have left school, but there are also those who have returned to alternative or regular programs. Some of these alternative programs are alternative routes to school completion (to a GED, for example) while others are intended to help students re-enter a diploma program. In addition, there are students who are in alternative programs to prevent dropping out, though they may never have left school. Finally, there are significant numbers of students who are chronic truants. There are many gradations of disengagement along the continuum between in-school status and dropout status. A fundamental choice made in the NELS:88 1FU was that any student who is receiving any kind of academic instruction -- whether that instruction is designed to lead to a high school diploma, a GED, or to neither -- should be administered the student questionnaire. Thus, students who were institutionalized (for example, in jail or reform school) completed the student questionnaire, as long as they received academic instruction, as too students in a home study situation (students who had left school and were being instructed at home owing to religious or other motives of their parents, or to disabilities), and those attending night classes at a school, church, or other setting. Only students who were receiving no academic instruction were administered the out of school (dropout) questionnaire. The situation of the NELS:88 1FU should be compared to that of HS&B.

A similar dropout definition was employed in High School and Beyond. The dropout definition employed in the HS&B First Follow-Up was "a person who was a high school sophomore in spring term of 1980 but who was neither enrolled in high school nor a high school graduate or the equivalent at the time of the follow-up survey in spring 1982."

More specifically, HS&B 1FU defined a dropout as a sample member who "has not attended school for the past month or more (not due to illness or accident), and does not intend to return."

Two dropout categories were provided as a further distinction within this general definition. One category encompassed dropouts neither enrolled in school nor attending a special program; the other embraced dropouts attending a special program (which might or might not be held within a school facility) such as a GED program or adult education courses, but which did not lead to a high school diploma per se. The reference period of "one month or more" is functionally equivalent to the twenty or more consecutive school days definition employed in the NELS:88 First Follow-Up.

In HS&B Base Year, alternative schools formed a separate sampling stratum. Various alternative programs were encountered in the course of the HS&B 1FU -- particularly continuation schools such as dropout centers, re-entry programs, pregnancy-maternity centers, evening and adult high schools, and street academies. (While we will confront the same kinds of programs in 1992, many of these programs, particularly in the area of dropout prevention, have of course grown in number and diversity in the years since HS&B 1FU.) The essential line of distinction enforced by HS&B was whether a particular alternative program led to regular high school completion, or to a GED, or to neither. Sample members were told that if they were enrolled in high school, they should complete the in school questionnaire. If they were attending a GED program or another special program, but were not enrolled in high school, they should complete the out of school questionnaire. Likewise, if they had left high school but were receiving no academic instruction, they should complete the out of school questionnaire.

One implication of the HS&B procedure that should be noted is that certain students who were considered dropouts in HS&B and received the out of school questionnaire -- for example, individuals pursuing a GED -- were considered to be students in NELS:88 1FU. Other participants in alternative programs -- say a program for pregnant students or students who are mothers, but still leading to a high school diploma -- were considered students in both NELS:88 1FU and HS&B. Since there were no "early dropouts" in HS&B, the fact that NELS:88 1FU dropout definitions do not readily map back into HS&B is not a serious difficulty. However, in NELS:88 2FU, whatever decisions are made about who should receive what questionnaire, the various "dropout" subpopulations must be distinguished in a manner that permits comparison of key NELS:88 statistics to those of HS&B.

### 4.1 Objectives and Content of the Second Follow-Up Field Test Dropout Questionnaire

**Objectives of the Dropout Questionnaire:** From the outset of the study, a principal design goal of NELS:88 has been to capture in a longitudinal dataset the entire subpopulation of school dropouts from within three distinct grade cohorts -- 1987-88 eighth graders, 1989-90 tenth graders, and 1991-92 twelfth graders. Early dropouts were identified in the First Follow-Up in 1990; most late dropouts will be captured in the Second Follow-Up, though the small number of students who leave school after a 1992 survey day will not be identified until the Third Follow-Up of NELS:88 in 1994.

It is important to note that the sample definition of a dropout and the intended respondent population for the dropout questionnaire are not strictly co-extensive. The event history definition of a dropout in NELS:88 permits temporary dropouts ("stopouts" -- that is individuals with a dropout episode who later return to school) to be classified as dropouts for sampling purposes (and thus followed with certainty and distinguished analytically). But the dropout questionnaire is intended for students who are currently out of school. (While it would be inappropriate to give an in-school student a research form intended for out-of-school students, there is some loss of information about the dropout experience of stopouts in consequence -- questions on reasons for dropping out in the first place; how the school, family and peers responded; and reasons for returning, would be particularly of interest to pose, were it feasible to do so.)

In addition to the general consideration of arriving at a proper mix of old (HS&B) and new (NELS:88) material, two special considerations were taken into account in designing the dropout instrument. These considerations were: (1) the need to speak to the diversity of situations encountered by early as contrasted with late dropouts and (2) the need to ensure comparability and articulation between student and dropout questionnaires so that comparisons can be made between school leavers and school completers on critical long-range outcome measures, all the while giving sufficient attention to the unique experiences of dropouts.

**Issue - Early dropouts, late dropouts:** NELS:88 offers the unprecedented opportunity to study, on a national scale with a large probability sample that includes significant representation of policy-relevant subgroups, both the early dropouts identified in the First Follow-Up, and later dropouts identified in the Second Follow-Up. HS&B dealt only with dropouts who left school after the spring of their sophomore year. From the point of view of simplicity and ease of questionnaire construction, the late starting point of HS&B was an advantage, in that the maximum out-of-school dropout spell to be covered in the HS&B First Follow-Up (twelfth grade, 1982) was less than two years. NELS:88 provides a more nearly unbiased sample of school leavers. A major complication in the construction of the NELS:88 2FU field test questionnaire, however, is that early dropouts who have not since returned to school will have been out of school for up to four years, while the more numerous late dropouts will have been out of school for two or fewer years. Questions about school experience will have a greater relevance and freshness for the late than for early dropouts. Indeed, the key school-related questions will already have been asked of the early school leavers, at the far more optimal time of the First Follow-Up. Ideally, one would wish to devote considerable questionnaire space to asking the more recent (late) dropouts about their past educational experiences. At the same time, one would wish to spend little or no time inquiring into the past school activities of the early dropouts; one would emphasize instead their current activities, plans for re-entering school, or participation in alternative workplace training or educational programs. Hence the dropout questionnaire was designed to accommodate both these groups by developing questions appropriate to the experience of early and late dropouts. In addition, the questionnaire was designed to maximize the quality of responses and the number of questions that could be asked by devising routing patterns that would ensure that the most relevant questions were posed to each group.

**Issue - Overlap and Articulation between In-School and Out-of-School Questionnaires:** A major NELS:88 design goal is to be able to compare the experience of in-school and out-of-school sample members; yet another goal is to be able to more fully explore the unique aspects of dropout dynamics and experience. These goals stand in a certain tension (they compete for the limited space) and complementarity (both goals must be realized if we are to achieve a complete picture) to each other. Hence a major objective of the instrument development process was to optimize the articulation -- both in form and content -- between the dropout and student questionnaires. In terms of content, this may be done by ensuring that key comparison measures are appropriately incorporated on each, while allotting sufficient space to exploration of the distinct out-of-school experience of dropouts. In terms of form, comparable question wording and format should be imposed, to the extent possible, though to some degree the generally more limited comprehension of school leavers, and differences of experience, militate against literal parallelism of item wording or response category format. The gap in reading and comprehension ability between dropouts and twelfth grade students cannot be eliminated, but its negative implications will be mitigated, if the recommendation of this report to further simplify questionnaire language and format is followed. To the extent that a gap remains, literal consistency must be sacrificed to consistency of meaning, a parallellism that may be achieved through further simplification, additional glosses, or other devices, aimed at improving the understanding of the instrument by dropouts.

**Content of the Dropout Questionnaire:** The NELS:88 2FU Dropout Questionnaire seeks to collect information that will furnish analysts with the opportunity to study both early and late dropouts on a national scale, to examine the contextual factors associated with dropping out, especially those related to school, and to create a profile of students in and out of school. The Dropout Questionnaire embraces six basic content areas:

I. **School experiences and activities.** Questions in this section elicit information about respondents' perceptions of and attitudes toward school and schooling, including teachers, homework, disciplinary rules, and the process of disengagement from school. This section will generate data for understanding the determinants of school completion and dropping out, the transition from school to the workforce and family formation, and effects of dropping out of school on self-esteem and future educational and career goals.

II. **Plans for the Future.** Data gathered in this section will sustain analyses of how education and career plans are linked to current behavior and to future achievement. These data will permit analysts to assess how changes in plans and aspirations relate to educational and out of school experiences that have occurred since eighth grade.

III. **Language use.** This section of the dropout questionnaire provides data about student and family language use patterns, and language ability, since 1990.

IV. **Self concept and attitudes.** This section gathers information about student self-esteem, locus of control, achievement orientation, life goals, and attitudes.

V. **Work experience**. This section identifies the types and amounts of work in which dropouts are engaged.

VI. **Family structure and environment**. The section on family structure and interactions inquires into household members, intergenerational closure, family decision making, and the home education support system.

### 4.2 Defining Dropout Sample Members

As described in Section 1.7, field interviewers identified 68 dropouts among the field test cohort members during the fall tracing activities. This classification was based on information obtained from the school and from the sample member or his or her parent/guardian.

Dropouts were also identified in the spring data collection phase during interviewers' pre-Survey Day phone calls with school coordinators. While confirming the enrollment status of 2FU field test sample members in these calls, school coordinators would inform field interviewers of students who had dropped out since fall tracing. As in the fall phase, interviewers were instructed to confirm the status of these students through the household with a parent/guardian or the sample member. Forty additional dropouts were identified during the spring data collection period, bringing the total number of dropouts in the 2FU field test to 108.

As experienced in the First Follow-Up, questions arose in the 2FU field test regarding the operationalization of the definition of a dropout. These ambiguities center around the status of sample members who are enrolled in non-diploma-granting programs (e.g., GED programs and non-academic programs that lead to certification). The issues involve both appropriate classification and questionnaire status. See Section 4.5 for a broader discussion of the issues and NORC's recommendation.

### 4.3 Group Administration Data Collection Procedures

Team leaders administered the Not Currently in School Questionnaire and the cognitive test to dropouts during off-campus group administration sessions. Team leaders were instructed to procure sites for these sessions that approximated as closely as possible the characteristics necessary for a Survey Day room; off-campus sessions were conducted in public libraries, community centers, and similar locations.

In off-campus survey sessions, team leaders followed the same procedures as for in-school sessions. Attendance was taken; permission was checked; in-school scripts and instructions were read; instruments were administered with the precise timing of an in-school session; and critical items were edited and retrieved.

Unlike in-school sessions, students who attended an off-campus session were reimbursed $15.00 for travel expenses at the end of the session. This reimbursement was not a payment for participation.

In few cases, it was preferable to administer the survey in a student's home. A home site off-campus administration was held when only one respondent in a particular area was eligible for an off-campus administration, the home environment was suitable, and a more desirable site was unavailable or inaccessible to the respondent. Team leaders followed the same procedures as for in-school and central site off-campus administrations. Respondents participating in home administrations did not receive the $15.00 reimbursement for travel expenses.

**Evaluation of Group Administration Data Collection Procedures:** Based on NELS:88 First Follow-Up experience, project staff were concerned about ensuring the attendance of dropouts at group data collection sessions. Since collecting data from dropouts in one-on-one situations had not proven to be problematic in the 1FU, a decision was made to focus 2FU field test resources on learning more about group administrations with this population.

A total of twenty-one group sessions were scheduled during the field period; 94 dropouts were invited to attend these scheduled sessions. By March 15, fifteen sessions had been held with only 20 dropouts completing the questionnaire and cognitive test. Field interviewers reported that several problems were contributing to the extremely low turnout: 1.) dropouts were generally working or engaged in other activities that severely restricted and complicated their schedules and 2.) the $15.00 travel reimbursement did not seem to be motivating dropouts to participate.

In an effort to learn how to get dropouts to participate in group administrations, field interviewers were authorized to recontact potential respondents and invite them to a session with the amount of reimbursement increased to $30.00 Six sessions were scheduled for March 18 through 23. An additional nine dropouts attended the sessions and completed a questionnaire and cognitive test. (See section 4.5 for recommendations for improving turnout to group administrations.)

Quality control procedures for the Dropout Questionnaire were very similar to those employed in Survey Day sessions. During the test administration, the team leader edited the Dropout Questionnaires, checking that critical items were completed in full. If data were missing, the team leader attempted retrieval at the sample member's work area when he or she had completed a test section. At the end of the testing session, sample members were instructed to close and hand in their test booklets. Any sample members with items yet unretrieved were asked to stay for a few minutes after the session.

### 4.4 Data Preparation

Data preparation encompasses all activities associated with converting information from the Dropout Questionnaire to data files. These included receipt, manual editing, coding and data entry.

**Receipt Control:** Team leaders were instructed to return Dropout Questionnaires and Cognitive Tests to NORC upon the completion of group off-site survey sessions. Overnight mail service was used to ensure the arrival of the questionnaire and test package and to expedite processing of the materials.

Dropout Questionnaires and tests were receipted in the Receipt Control Shop of NORC's Lake Park Facility. Receipt control clerks updated dispositions in the field test Survey Management System (SMS). Separate SMS variables were used to track questionnaire and test status. The current questionnaire and test dispositions were entered directly from the completed instruments, in a batch update mode by type of document and disposition. As a final step, the main disposition for each case was updated. Questionnaire and test dispositions were again updated when questionnaires were sent to data entry and tests were sent to ETS.

**Editing/Coding:** Dropout Questionnaires were edited for missing items in the Edit Shop of the NORC Lake Park Facility. The coding system used was consistent with that from the First Follow-Up Field Test. This edit consisted primarily of application of standard reserve codes. Marginal indications of "Refused" or "Don't know" were coded as "7" and "8", respectively. Unmarked missing items, whether skipped legitimately or illegitimately, were marked "9" for missing, and a "6" was used to indicate an illegitimate multiple response. Open-ended questions were coded "1", to indicate that data were present. These reserve codes were preceded by a "9" for some items, depending on the number of columns corresponding to the question. For example, a missing item with response codes ranging from 00 to 12, with a marginal indication of refusal, was coded a "97".

**Data Entry:** Questionnaires were data entered at our Lake Park Facility using a computer-assisted date entry (CADE) software package. Critical items were 100 percent verified (that is, re-entered independently by a different operator), with the software comparing all fields for exact matches. Given the purposes of the 2FU Field Test -- to facilitate the identification of question logic and skip problems within the questionnaire -- no machine cleaning of the data was performed.

### 4.5 Recommendations for Changes in Dropout Data Collection Procedures

Because obtaining data about dropouts is a major policy concern of the NELS:88 project and of NCES, special care will be taken in the 2FU main study to identify, trace, and collect data from dropouts. Experiences in the First Follow-Up and the Second Follow-Up Field Test lead us to make recommendations for changes in the procedures used to collect data from this important group.

**Definition of a Dropout:** Several ambiguities have arisen in the 1FU and 2FU field test concerning the definition of a dropout. While the traditional "twenty consecutive days" guideline seems straightforward, our contacts with these sample members constantly present new questions and diversities.

In considering this question, it is important to note that there are three dimensions to the definition issue. First is sampling status. Dropouts -- and stopouts, or sample members who were previously found to be dropouts but are currently back in school -- are sampled with certainty; student sample members are not. The second dimension concerns which questionnaire the sample member should be administered and, within that questionnaire, on which school he or she should be asked to report. Third, it must be remembered that many young people who are enrolled in nontraditional instructional programs consider themselves to be in school.

With regard to the sampling status portion of the ambiguity, the primary question has been how NELS:88 defines "in school". With the diversity of programs for dropout prevention, high school completion and equivalency certification, and certification in vocational trades, many questions come to us from our field staff regarding the classification status of individual sample members. For the selection of students and schools for the main study sample, we are currently defining "in school" as attending an instructional program that can lead to the attainment of a high school diploma. Enrollment in any other instructional programs (GED programs, other equivalency certifications, vocational certification, reform school, prison educational program, etc.) results in the sample member's classification as a dropout or alternative completer. No decision has been reached about the sampling status (certainty or noncertainty) of these cohort members; cost constraints may prohibit their inclusion with certainty.

The second and third issues -- administration of which questionnaire and the self-identification of the sample member -- are interwoven. As indicated in Section 5.2.3, dropouts are greatly confused if one inquires vaguely about "school"; for many, this can mean the last high school they attend or some other type of school attended since their departure from the diploma-granting institution. For example, in the 1FU any sample members responding "yes" to the question "Are you currently enrolled in school?" were administered the Student Questionnaire. It is clear, however, that what constitutes a "school" in one circumstance may not be so defined in another.

We recommend screening sample members with a dropout status (and others, as the need arises) before data collection, much as was done in HS&B to gather information about the context of the situation. (See HS&B form on next page.) The screener could be administered by telephone in the interviewer's attempt to schedule an appointment for data collection, thereby containing costs.

Alternatively, this screening process could be accomplished through the fall interaction with school staff rather than with the sample member, but this would increase the level of effort in this visit and the level of burden on school staff. If clarity of data is the only significant issue, a question can be added to the Student Questionnaire asking the sample member if the program he or she is enrolled in can culminate in the award of a high school diploma or some other sort of certification, with follow-up subitems to obtain the specific context of non-diploma granting programs.

The remaining issue is defining which sample members are to be administered the Dropout Questionnaire, which the Student Questionnaire, and in both situations on which school the sample member is to report in completing the instrument. There is no clear precedent to follow. As indicated on the HS&B form, those sample members attending instructional programs to earn a GED and other special programs were administered the Dropout Questionnaire. In the NELS:88 1FU, any sample member reporting "enrollment in school" was administered the Student Questionnaire as described above. Both have advantages and disadvantages. We recommend that this issue be disccused and a consensus decision reached for implementation in the NELS:88 2FU main study.

**Group administration:** Plans for the NELS:88 Second Follow-Up main study include data collection from a majority of dropouts in group sessions. Our experience in the 1FU and 2FU field test suggests that ensuring such a large proportion of dropouts to attend group administrations will be more difficult than experienced in HS&B.

To devise an effective plan of action, the components of the group administration problem must be clearly understood. Based on anecdotal information from the 2FU field staff, it seems apparent that the problem is multi-faceted. Following is a summary of some of the obstacles we face.

a.) Dropouts are a diverse group and often lead lives more complicated than those of other sample members in the cohort. Interviewers reported that a significant number of dropouts they contacted are married, have a child or have a full-time job or multiple part-time jobs that often include evening and weekend hours.

b.) While dropouts are generally disproportionately of lower SES, the $15.00 travel reimbursement that proved successful in HS&B group administrations seems not to be as attractive in 1991 as it was in 1982. A dropout employed in a minimum-wage position can earn almost as much in three hours as we are currently authorized to offer him or her for the same amount of time.

c.) In that three-hour block of time, we are requesting that the dropout complete a self-administered questionnaire of intimidating length and a cognitive test battery. We suspect that the 1FU data will suggest that dropouts have more limited reading skills and performed more poorly in school before dropping out than their in-school counterparts. If this is indeed true, we are asking these sample members to engage in activities that are more difficult, frustrating and demoralizing. And we require them to do so not in the privacy of a one-on-one situation, but rather in a group setting with others their own age (some of whom will be early graduates).

## HIGH SCHOOL AND BEYOND

### GROUP ADMINISTRATION SCREENER

NAME ______________________________________

Please circle the number of the statement that fits your present school status. Please read each statement before you choose one. If none of the statements applies to you, or if you have any questions about them, please ask the Survey Representative.

| | | OFFICE USE |
|---|---|---|
| I am now attending the same high school that I attended when I was a sophomore in the spring of 1980 | 1 | SO/Q |
| I am not now enrolled in the school I attended when I was a sophomore in the spring of 1980, because -- | | |
| I am currently enrolled in another high school | 2 | TR/Q and SO/Q |
| I graduated from high school before March 1, 1982 | 3 | EG/Q and SO/Q |
| I am attending a GED program, or another special program, but am not enrolled in high school | 4 | DO/Q |
| I have left high school; that is, I haven't attended school for the past month, or more | 5 | DO/Q |
| None of these (Please explain) | 6 | REVIEW |

Please return this form to the Survey Representative. Thank you.

Based on the importance of collecting as much data as possible from these sample members, we propose the following changes to procedures for the 2FU main study data collection. The focus of these recommendations is to maximize response rates for this difficult population while containing the costs of data collection to the greatest extent possible.

**Mode of Administration:** Main study plans should be altered to reflect the difficulty of ensuring dropouts attend group sessions. Conversely, plans will include more one-on-one administrations. We also anticipate successfully collecting questionnaire data from a small percentage (no more than 15 percent) of dropouts by mail and/or telephone interviews. This will allow interviewers to be more flexible and react, within broad guidelines, to the peculiarities of individual situations. For example, if a hard-to-locate dropout is finally found, the interviewer can use her or his discretion in scheduling a one-on-one or group session, or collect the data immediately.

**Increased Incentive:** We recommend that interviewers be authorized to offer an increased incentive to potential respondents. Rather than only offering a fixed fee, interviewers should be trained to be sensitive to each potential respondent's particular situation and to offer a "package" incentive (not to exceed $50.00) that will be most attractive to the sample member. For example, if the potential respondent is a single mother who depends on mass transportation, the interviewer would offer a reasonable amount to reimburse her for child care and transportation costs in addition to a set incentive of $30.00 This $30.00 fee can be described as $10.00 per hour for the session. Incentive payments should be distributed to the completers in envelopes to minimize the possibility of the variable fee becoming a source of argument and hard feelings. The receipt signed by the respondent would specify the components of the total fee (e.g. $5.00 for transportation, $15.00 for child care, $30.00 for completing questionnaire and test).

**Reminder Postcard:** In addition to the permission form letter, potential dropout respondents should be sent a postcard reminder similar to the Survey Day invitation distributed to in-school sample members. An alternative approach already employed by many interviewers is to call each dropout scheduled to attend the session as a reminder of their commitment and the location and time of the session.

**Group Site Rental:** In the 2FU field test, we experienced a surprising reluctance to allow NELS:88 the use of private and public space at no charge. This is perhaps due to concerns about legal liability and potential law suits if an accident happens. To overcome this minor obstacle, we recommend increasing the funds available for interviewers to rent sites appropriate for a group administration. Sites that worked best in the field test included community centers, recreation centers, private rooms in casual restaurants, and small hotel meeting rooms.

**Additional Training:** In combination with simplifying the questionnaire wording and format, we recommend that interviewers be trained to recognize when respondents are having difficulty with the questionnaire and to properly provide these participants with more individual assistance, as it is needed, to overcome deficiencies in reading skills.

## Chapter 5: Dropout Survey Data Analysis

This chapter discusses the results of an in-depth analysis of qualitative and quantitative data collected using the Dropout Questionnaire. Both general and specific recommendations are offered for item refinement, in the interest of collecting high quality data in the main study.

The quantitative data consist of the responses of 29 dropout sample members to whom the instrument was administered in group off-site sessions. A standard item nonresponse analysis was performed on the data. Because of the small number of field test dropouts surveyed, First Follow-Up Main Study Dropout Questionnaire frequencies were also examined, when item content between the First and Second Follow-Up instruments overlapped. To supplement these data sets, NORC professional staff also collected qualitative data for use in assessing the questionnaire, employing the technique of cognitive interviewing. Findings from the cognitive interviews form the foundation of the majority of recommendations made in this section because of the detail of the information provided by the cognitive interviewing technique and the small number of pretest observations collected in the field.

### 5.1 Dropout Questionnaire Response Rates

Overall, 29 out of 108 dropout sample members with whom data collection was attempted participated in the study, for an overall response rate to the Dropout Questionnaire of 27 percent. A much higher response for the dropout survey could have been attained. The deliberate decision to not attempt to achieve a high response rate was made for several reasons. First, our experience in the NELS:88 First Follow-Up suggested that data collection procedures for personal interviews with dropouts are satisfactory and did not need further assessment. (Over 91 percent of sample members who had dropped out were successfully surveyed in the First Follow-Up.) Second, the main purpose of the 2FU Field Test of the dropout survey was to find effective methods for surveying dropouts at group/off-site sessions in the main study. It was determined that these procedures could be adequately tested on a relatively small number of dropout respondents. Third, to so limit the scope of the 2FU Field Test dropout survey resulted in cost containment. For these reasons, initial refusals and survey session no-shows were not vigorously pursued as they will be in the main study.

### 5.2 Analysis of the Dropout Questionnaire

As stated above, the primary goal of the NELS:88 Second Follow-Up Field Test dropout component was to test, assess, and refine procedures for group data collection sessions. By pursuing that objective, completion of the Dropout Questionnaire and cognitive test battery by field test sample members was artificially minimized. However, even if all available dropouts had been surveyed, the ability to make generalizations like those presented for the Student and Parent Questionnaires is limited by a rather small sample (108 dropouts were identified from the longitudinal cohort and augmentation sample). To supplement the data collected through these standard procedures, field test staff decided to collect and analyze qualitative data for use in assessing the questionnaire. The cognitive interview technique was employed to obtain these qualitative data.

This section describes the cognitive interview methodology, its application to this component of the NELS:88 2FU Field Test, and the individual problems and recommended solutions identified through this process. Data obtained from the interviewer-administered group sessions and from the First Follow-Up main study dropout questionnaire are cited in the last section in situations in which such data enlightens the investigation. The 1FU data are especially suitable for qualitative assessment because of the large number of cases (n = 867), although it must be remembered that respondents in the 1FU are two years younger than 2FU respondents will be.

### 5.2.1 The Cognitive Interview

The cognitive interview was developed by researchers in the interdisciplinary field of cognitive aspects of survey methodology. Several NORC staff -- notably Norman Bradburn, Roger Tourangeau, David Mingay, and Kenneth Rasinski -- have been pioneers in the developing field. The practical value of cognitive interview procedures is evidenced by the fact that, although the procedure was introduced only five years ago, three federal statistical agencies now routinely use cognitive interviews to develop and revise their survey instruments. (Those agencies are the U.S. Bureau of the Census, the U.S. Bureau of Labor Statistics, and the National Center for Health Statistics.)

Survey researchers have long known that respondent error is often a significant threat to the quality of survey data. Causes of this type of error include the use of complex or ambiguous words or phrases, asking for information that the respondents are unable or unwilling to report, and requiring that respondents make judgments and estimations that are beyond their knowledge or ability. Field pretesting provides adequate information on some aspects of a survey such as question flow, skip patterns, respondent burden, and field procedures. However, field testing does little to provide information on the respondent's understanding of specific words and terms or on respondents' recall and response processes. In fact, traditional field testing may produce data which seem quite reasonable, with almost all responses to an item falling within an expected range, but with no indication that the responses are invalid. In marked contrast, the cognitive interview not only identifies problems of the type described above, but provides sufficiently detailed information about why questions are being poorly answered so that improvements can be readily made.

The conceptual framework underlying the cognitive interview technique conceives of responding to a survey question as involving a sequence of four cognitive tasks. First, the respondent must comprehend the meaning of the question. Second, the respondent must decide what information should be retrieved from memory and how to search their memory to retrieve it. Third, when relevant information has been retrieved, the respondent engages in an estimation or judgment process in which he or she evaluates the information retrieved and its relevance to his or her interpretation of the question. The respondent may combine several pieces of information to form a response or may decide to use the retrieved information as a starting point in forming an estimated response. Finally, the respondent weighs factors such as the sensitivity of the question, the social desirability of the answer, and the probable accuracy of the answer. A response is then decided upon and provided (Jobe and Mingay, 1986). At any of these stages the respondent may decide that she or he has made an error, or that there is a better way to obtain the answer, and return to one of the previous cognitive stages. In addition, a respondent who is not motivated to give a carefully thought out response may omit some of these stages or apply them in a cursory way in order to be able to answer quickly and with a minimum of effort (Krosnick and Alwin, 1987).

Cognitive interviewers employ a variety of specific methods, depending on the issues being investigated. A few of these are described below.

- **Concurrent think-aloud interviews:** These exploratory interviews are relatively unstructured. The respondent is asked to read the questions for the interviewer and to think aloud as he or she answers the questions. This technique is especially useful for studying recall and estimation processes. The interviews are usually audio taped, so the interviewer can concentrate on probing the responses and can analyze their content at a later time.

- **Paraphrasing:** The respondent is asked to repeat the question in his or her own words. This can be used effectively to determine if a question contains words or phrases not readily understood and whether a question is so complex that the respondent is missing important qualifiers (such as a reference period).

- **Confidence ratings:** After answering selected questions, the respondent is asked to rate the degree of confidence she or he has in the accuracy of their answers. This may indicate to what extent the respondent had difficulty formulating an answer to a question and whether he/she was guessing. Although there are often other clues the respondent is having difficulty, this is a method for systematically assessing the extent of the problem (Roysten et al., 1986).

Most of these methods are more time-consuming than the usual interview, so it is often necessary to study a questionnaire one section at a time or to target certain items, to avoid excessive respondent fatigue. Cognitive interviewing is often employed in several iterations -- problems are identified in a series of cognitive interviews, changes are made to correct the problems identified, and the revised items are tested in another series of cognitive interviews.

### 5.2.2 Use of Cognitive Interviews to Assess the NELS:88 Second Follow-Up Field Test Dropout Questionnaire

Preparations for conducting cognitive interviews on the NELS:88 2FU Field Test Dropout Questionnaire began as data collection activities slowed in mid March. One of the first preparatory activities was the selection of items to be tested. Given the length of the instrument, cognitive interviews could not be conducted on the entire questionnaire. Items were identified that had been problematic in the past (either in prior rounds of NELS:88 or in HS&B), were designated as critical (that is, of particular importance to analysts), or were considered to have difficult wording or format by the field test staff. The items selected for cognitive interview testing, as well as interviewers' comments from the completed sessions, appear in Appendix D.

Recruiting methods were designed to obtain a diverse group of respondents, with particular attention to targeting those respondents anticipated to have the most trouble with the questionnaire (e.g., poor readers, early dropouts). Respondents were recruited through contacts with and referrals from administrators of community service agencies in Chicago and Washington, D.C., through flyers, and through networks maintained by NORC professional staff. Respondents were volunteers and may differ in important respects from those individuals who do not volunteer. However, probability sampling recruitment procedures are usually unsatisfactory for recruiting individuals to participate in cognitive interviews because of time and cost constraints and because it is essential that respondents be highly motivated and willing to work hard at reporting how they are arriving at their answers.

Ten cognitive interviews were completed -- five with high school dropouts from the Chicago area and five from the Washington, D.C. area. Of these ten, eight were African-Americans, one was Hispanic, and one was Thai. Eight respondents were female, two male. All were between the ages of 18 and 20 years old. Interviews were conducted by NORC professional staff trained in cognitive interviewing. Specific training for this project was conducted by David Mingay and Roger Tourangeau (both of whom served as cognitive interviewers), in consultation with field test staff.

### 5.2.3 Qualitative and Quantitative Analysis of Dropout Questionnaire Data

In this section we discuss items found to be problematic in the Dropout Questionnaire, as identified by analysis of the results of the ten cognitive interviews conducted with dropouts in Chicago and Washington, D.C., item nonresponse analysis of the 29 Dropout Questionnaires completed during the field test, and review of the First Follow-Up Dropout Questionnaire frequencies. Our primary focus is on the qualitative data from the cognitive interviews, which are summarized in Appendix D. In analyzing the quantitative data, an item with a nonresponse rate greater than ten percent was considered to be problematic. A threshold of ten percent, rather than the 20 percent used in the student item nonresponse analysis, was used because of the greater attention dropouts received in the smaller group off-site sessions.

In our analysis we discovered that item nonresponse and respondents' difficulties with items on the Dropout Questionnaire have three major sources: (1) poor question design; (2) difficult, vague, or abstract language; and (3) question overcrowding.

Question 17 is a particularly salient example of a poorly designed question. It attempts to collect too much information in one question; it seeks to determine whether the respondent has a GED or equivalent, whether he or she plans to go back to school to obtain a diploma, whether he or she intends to earn a GED or equivalent, and whether or not he or she has enrolled in a GED preparatory class.

Question 25 provides a good example of how imprecise wording can lead to respondents misinterpreting the intent of a question; respondents interpreted "your high school program" to mean the general program orientation of the school, rather than the type of courses in which the respondent was enrolled. Several of the explanation boxes also use difficult or abstract language, and, rather than facilitating responding, confuse the respondent.

Question 60, which inquires about participation in job training programs or courses, amply illustrates question overcrowding. The format is so crowded and confusing that respondents find it difficult to tell which parts of the question they should complete.

In addition to these three potential sources of error or nonresponse, another problem, of a mechanical rather than substantive nature, is respondents' circling the response option rather than the response code. The cognitive interview results and a review of the completed field test questionnaires suggest that this may not be a rare occurrence, particularly with "Yes/No" questions. In circling the words rather than the code, respondents may miss skip instructions. We suggest that team leaders strongly emphasize to dropouts the importance of recording responses correctly.

Respondents experienced other difficulties with items on the Dropout Questionnaire which are not easily categorized. A discussion of problematic items on the questionnaire, along with recommendations for their revision, follows.

**Q17 Do you plan to get a high school diploma or GED?**

This question was particularly vexing for five of the ten cognitive interview respondents. One respondent had difficulty following the skip instructions, because of the poor formatting of the instructions for categories "02" through "04". Three of the five found choosing an appropriate category difficult. One suggested that the initial "No" or "Yes" might confuse respondents and cause them to misread the response options. Another, who ultimately chose option "04" --"Yes, I plan to take an equivalency test such as the GED, but I have not yet enrolled in a class, or don't plan to enroll" -- attributed her initial indecision to not reading the response categories thoroughly and consequently missing the applicable phrase -- "don't plan to enroll" -- in category "04". The third respondent felt that none of the options accurately described his situation, since he had been enrolled in a GED preparatory class, but was not currently, and neither "03" nor "04" seemed appropriate.

**Recommendation**: We recommend that item 17 be divided into three questions, to simplify the skip pattern and to facilitate response selection, as follows.

17A. Do you plan to get a high school diploma or a GED?

|  | (CIRCLE ONE) |  |
|---|---|---|
| No | 1 | SKIP TO QUESTION 20 |
| Yes | 2 | GO TO QUESTION 17B |
| I already have a GED or equivalent | 3 | SKIP TO QUESTION 19 |

17B. Are you currently enrolled in a preparatory class for taking the GED?

|  |  |  |
|---|---|---|
| No | 1 | GO TO QUESTION 17C |
| Yes | 2 | SKIP TO QUESTION 18 |

17C. Do you plan to...

(CIRCLE ONE ON EACH LINE)

|  | No | Yes |
|---|---|---|
| Go back to school to get a diploma or GED? | 1 | 2 |
| Enroll in a class to prepare for taking the GED? | 1 | 2 |

**Q18 When do you expect to receive a high school diploma, GED, or High School Equivalency certificate?**

All ten cognitive interview respondents were eligible to answer Q18, and only one answered the question easily and followed the skip instructions correctly. The other respondents had difficulty answering this question because they either did not know when they would receive their diploma, GED, or equivalency certificate and would not hazard a guess without probing, or knew when they expected to take the GED exam but did not know when they would receive the actual degree.

**Recommendation:** A "Don't know" response option, appearing below the month and year spaces, should be added to the question. If the retention of the question stem in its present form is desirable, we recommend that "about" be inserted before the initial "when", to indicate more explicitly the acceptability of an estimate. However, it may be desirable to revise the question so that it inquires about the date the respondent expects to take the GED or High School Equivalency exam. A suggested rewording is "Approximately when do you expect to receive a high school diploma, or to take the exam for the GED or High School Equivalency certificate?"

**Q20 Since leaving school, have you enrolled in an educational institution, such as vocational or trade school, or college?**

The present skip instructions in Q17 and Q18 route only respondents who have a high school diploma or GED to this question. However, one of the cognitive interview respondents who didn't have a diploma or GED did attend a trade school after leaving high school.

**Recommendation:** Though it is difficult to make a recommendation based on one case, it does seem probable that some dropouts in the main study will have attended a postsecondary school without having obtained a high school diploma or GED. It may be desirable to have all respondents answer this question. If so, Q20 should explicitly refer to leaving high school.

**Q23 How much do you agree with the following statements about the school you left?**

First Follow-Up main study nonresponse rates for the subitems of this question were high, ranging from 11.6 percent to 12.3 percent. The high nonresponse to these items may be attributable to the number of subitems (fifteen in the First Follow-Up main study version, seventeen in the 2FU field test version), respondents' lack of interest in the subject (school), and recall problems.

**Recommendation:** We recommend either that the question be divided into two questions, to reduce apparent burden and increase response, or that some subitems be eliminated. We welcome suggestions for subitems to be deleted.

**Q25 In the last high school you attended, which of the following best describes your high school program?**

All but one cognitive interview respondent had difficulty answering this question, for several reasons. The question stem is ambiguous. Probing revealed that some respondents interpreted "your high school program" to mean the prevailing instructional orientation of the last high school they attended, rather than the program in which they were enrolled. Others did not understand what "high school program" means. Some respondents either did not understand the response categories or appeared not to read them thoroughly before answering.

Of the 29 field test respondents, four did not answer this question, a higher nonresponse rate than for surrounding items.

**Recommendation:** The ambiguity in the question stem should be clarified and "high school program" should be defined. An alternate wording might be: "In the last high school you attended, which of the following best describes the type of courses you took?" In addition, collapsing some of the vocational, technical, and business and career categories may be seemly, since dropouts, unlike school principals, may not be able to make the required distinctions between these categories.

**Alternative Program Explanation Box**
**Q26 Have you ever participated in an alternative program?**

Seven of the nine cognitive interview respondents who read the explanation box still did not understand the meaning of "alternative program". Obviously, this affected how they answered the following item, Q26.

**Recommendation:** We suggest that the wording of the explanation be simplified, and the meaning of "alternative program" be clarified further, possibly through the enumeration of examples of alternative programs. The revised explanation might read: "The next few questions are about alternative programs. An alternative program is a special program that helps students stay in or return for completion of high school or an equivalent like the GED. An alternative program may be part of a regular high school or may be separate from any school."

**Q27 When did you enter the most recent alternative program?**

Of the nine field test respondents who were eligible to answer this question (based on their response to Q26), four did not report the month in which they entered the most recent alternative program; all, however, did report the year in which they did. A similar nonresponse pattern was found in Q11, Q18, Q57E and other questions requiring the reporting of month and year. Results of the cognitive interviews also indicate that, in answering such questions, respondents have difficulty recalling the month in which certain events occurred.

**Recommendation:** We suggest that the response format used in the First Follow-Up Dropout Questionnaire, which listed months and years, be used for this and similar items in the Second Follow-Up instrument.

**Q31 Did you receive any of the following services from this program, or were they not offered?**

One of the cognitive interview respondents interpreted the final response column, "Was not offered", to mean "was not offered to me". This ambiguity should be clarified. The word "peer" in subitem c, "Peer tutoring", also was not understood by at least one respondent, and its use is likely to present difficulties in the main study.

**Recommendation:** Presumably, "Was not offered" means that the program did not offer the service. We recommend that the response column heading be changed to "Was not offered by program". We also recommend that "Peer tutoring" be supplanted by "Tutoring by other students".

**Q37 How often do you spend time on the following activities?**

First Follow-Up nonresponse rates for the sixteen subitems of this question are high, ranging from 13.7 percent to 15.2 percent. The high nonresponse probably can be attributed to the sheer number of subitems; the burden associated with asking the respondent to estimate the average number of times a week he or she participates in certain activities may also contribute to nonresponse.

**Recommendation:** We suggest that the following subitems be eliminated from this question: E, F, J, K, and O. Our rationale for deleting these items is that they promise to be less analytically useful than the other subitems of the question in identifying correlates of participation in community, family, peer-group, and individual activities.

**Q39 How far in school do you think your father and your mother want you to go?**

Of the 29 Field Test respondents, 11 gave invalid responses to this question. The majority of this nonresponse was due to multiple responses; some respondents circled as many as seven response options. Many of these multiple responses are nonsensical. One respondents selected "Less than high school graduation", "Graduate from high school, but not go any further", "Attend a two-year college", "Graduate from college", "Attend a higher level of school after graduating from college", and "Don't know". The Student Questionnaire version of this item did not demonstrate high nonresponse. Dropouts may have more difficulty understanding the instructions to circle only one option and the implicit logic that one should not circle both "Attend a four-year college" and "Graduate from college". The educational aspirations dropouts' parents have for their children may also be less focused, and this may surface in the confused multiple responses given by dropouts. Similar nonresponse, or multiple responding, is not found in Q40, which asks about the respondent's educational aspirations, because the question is a critical item and therefore subject to field scan editing and retrieval.

**Recommendation:** We recommend that the instructions, "CIRCLE ONE OPTION", be made bolder. Response to this question may also be improved by dividing it into two questions, which ask about paternal and maternal aspirations separately. The crowdedness of the present two-column format may attenuate respondents' concentration and blunt the impact of the instructions.

**Q40 As things stand now, how far in school do you think you will get?**

The majority of cognitive interview respondents had difficulty answering this question. One did not know which response option to select, since he expects to obtain only a GED. Others may have been confused by the crowding of the response options. Use of the word "school" in the question stem may also have confused some respondents, since they might have initially thought that the question did not apply to them. Dropouts also probably have difficulty projecting how far in school they will get.

**Recommendation:** We recommend the addition of a response option for GED or other high school equivalency degree, even though such an option may not be in strict accord with the question stem. It may be desirable to eliminate "school" from the stem, so that the question reads: "As things stand now, how far in your education do you think you will get?" A blank line should also be inserted between the response options; the "College program" response options appear particularly crowded.

**Q42 Job or occupation at age 30** (Write-in)
**Q42A Category best describing job in Question 42**

The majority of the cognitive interview respondents had no problem answering Q42. However, four of the ten respondents miscoded the jobs or occupations that they reported in Q42 in answering Q42A.

Nonresponse to these questions was relatively high among the field test respondents. Five of the 29 failed to answer Q42, and five also did not answer Q42A.

**Recommendation:** The changes recommended for the Student Questionnaire version of the occupational coding item should also be implemented for Dropout Questionnaire item Q42A. (See Section 3.9)

**Q43 How much education do you think you need to have in order to get a job in the area that you chose in the last question?**

Nonresponse to this question was relatively high among the field test dropouts. Of the 29, five did not respond to this item. Four of the five nonresponses were multiple responses.

**Recommendation:** Again, emphasizing the "CIRCLE ONE" instruction may reduce multiple responding. It may also be desirable to add further instructions, explaining that one should not, for example, circle both "High school diploma" and "College degree".

**"Money and Work" Explanation Box**

Three of the seven cognitive interview respondents did not understand the explanation of the Money and Work section's purpose. In fact, rather than focusing their concentration, it appeared to confuse them.

**Recommendation:** The explanation should be simplified. Phrases such as "the relation of your training and education to your work experience" and "reasons for changes in job situations" probably add nothing to respondents' understanding of the section's purpose. A better explanation might be:

> This section asks about the kinds of jobs you have had, your work hours, and your income from these jobs. We are also interested in how satisfied you have been with your jobs and in the impact your education and training have had on your work life.

**Q55 Why weren't you looking for a job last week?**

Multiple responding among the field test dropouts was again a problem for this item. Three of the four nonresponders selected more than one category. Apparently they had more than one reason for not looking for a job in the previous week and disregarded the instruction to circle only one response option.

**Recommendation:** We recommend that the sentence "IF YOU HAD MORE THAN ONE REASON FOR NOT LOOKING FOR A JOB LAST WEEK, CIRCLE THE MOST IMPORTANT REASON" be added below the question stem.

**Q56 How many jobs have you held since you last left school?**

The findings of some of the cognitive interviews suggest that two ambiguities in this question may confuse respondents and adversely affect data quality. Some respondents may include occasional jobs, such as babysitting, street vending, and hairdressing, when reporting the number of jobs they have held. The school to which the question refers is also ambiguous, when the respondent has attended a postsecondary school since leaving high school.

**Recommendation:** We recommend that the question stem be modified to: "How many regular (part-time or full-time) jobs have you held since you left high school?"

**Q57 CURRENT OR MOST RECENT JOB HELD SINCE YOU LEFT SCHOOL. IF YOU HAD TWO JOBS AT THE SAME TIME, ANSWER FOR THE JOB YOU HAD THE LONGEST.**

Four of the six cognitive interview respondents were confused by these instructions. Either they did not understand the difference between "current" and "most recent", or they focused on the second sentence rather than the first. Some respondents were also troubled by the use of both past and present tenses in the stems of several of the following items (Q57A, Q57I, Q57J, Q57K). Clearly, the attempt to collect job information from those who are currently employed and those who are not in the same question series confused respondents.

**Recommendation:** To clarify the job of reference and to simplify skip patterns slightly, we recommend that this series of questions be preceded by a filter question, to route respondents who are not currently employed and those who are through different sets of questions. The filter would simply ask "Are you currently employed?" Dropouts responding "Yes" would be routed through a series of items consisting of Q57A through Q57E and Q57I through Q57O. The second set of items, for respondents not currently employed, would comprise Q57A through Q57O; the job of reference for these items would be the respondent's last job. Verb tenses would, of course, be appropriate to the employment status of the respondent.

More generally, we strongly recommend further qualitative assessment of Q57 and Q58. Given the importance of data on the employment experiences of dropouts, we urge further efforts to simplify the instruments for this series of items and to refine question wording.

**Q57E When did you start looking for this job?**

Two cognitive interview respondents reported that they could not answer this question, because they had not looked for their present job; it was offered to them.

**Recommendation:** Add a "Does Not Apply" response option.

**Q57G Why did you leave this job?**

One cognitive interview respondent reported that she left her job because it was seasonal work. Rather than select the first response option, "Job ended", she circled the "Other" category.

**Recommendation:** Though only one of the ten respondents had difficulty answering this question, it is advisable to add "seasonal" to the examples following "Job ended" ("temporary job, laid off, or fired").

**Q57M In this job are/were you...**

This question asks the respondent to categorize his or her employer. All but one of the cognitive interview respondents had difficulty selecting a response option. From their comments, it appears that they did not understand the response options, particularly the first, "An employee of a **PRIVATE COMPANY**". One of the dropouts did not know that a drug store and a restaurant are both private companies; apparently, most of the others also did not understand what constitutes a private company.

**Recommendation:** We suggest that examples of the various types of employers, particularly examples of private companies, be added to the response categories. However, some respondents may erroneously not select the response category if the specific field of their employer (restaurant, drug store, temporary agency) is not listed among the examples. Another option is deleting the word "private". Both alternatives should be tested in subsequent cognitive interviews to identify the optimal solution. Also, the ambiguity in the fourth and fifth options, "Working **FOR/WITHOUT PAY** in family business or farm", should be clarified. "Family business or farm" should be modified to "your family's business or farm."

**Q59 How much of the money you make is spent on each of the categories listed below?**

One cognitive interview respondent assumed that this question referred to the first job he had after leaving high school, the subject of Q58, rather than his present job because of the poor transition between questions.

First Follow-Up main study nonresponse rates for the subitems of this question are high, from 13.1 percent to 15.2 percent. This nonresponse may be due to respondents skipping items for which they should have circled the first response option, "None of it".

**Recommendation:** We recommend that the question refer explicitly to the respondent's current job.

**Q60 Have you participated in and completed any of the following types of training programs or courses?**

Cognitive interview respondents were very confused by this question. They did not understand the response options. Respondents did not know what a "Formal Registered Apprenticeship" is, what constitutes a "Government training program", or the meaning of "Correspondence course". One respondent had difficulty with the term "participated". They were also bewildered by the admittedly confusing format of the question.

**Recommendation:** The question stem should more explicitly state that it refers to **job** training programs or courses. "Correspondence course(s)" could easily be changed to "Course(s) by mail or television". Examples of government training programs also could be given, and the response option should specify that "government" includes federal, state, and local. "Union apprenticeship" may be more intelligible than "Formal Registered Apprenticeship". Presumably, if a respondent is enrolled in such an apprenticeship, he or she will recognize it and select this response option. We further recommend that the question stem be modified to eliminate its double-barrelled nature (participated in and completed). If measuring participation is the primary research objective of the item, we suggest that "attended" replace "participated in".

The question should be reformatted, so that respondents can more easily wend their way through it. It may be desirable to split the question into four distinct questions.

**Q69 How do you feel about each of the following statements?**

**Q70 Choose the answer that is best for you.**

**Q71 Think about how you see your future. What are the chances that...**

First Follow-Up nonresponse rates for the subitems of these attitudinal questions are high, ranging from 13.0 percent to 14.6 percent. This nonresponse can be attributed to the length of the subitem lists, the personal nature of the questions and sensitive nature of some of the items, and the burden associated with the response formats.

Field Test nonresponse to these items was nonexistent or very low.

**Recommendation:** Retain items in their present form. The level of nonresponse to these items in the First Follow-Up is acceptable (given their nature), and response may increase in the Second Follow-Up with an older sample. It may be helpful to repeat the response categories midway through Q70 and Q71. Elimination of subitems probably is not desirable.

Q76 **At what age do you expect to...**

For items Q76A, Q76B, and Q76E, field test dropout nonresponse was relatively high. Four respondents did not answer the first item, and three did not answer either of the last two. Respondents might not have answered these items, which ask about the age at which the respondent expects to marry, have his or her first child, and finish his or her full-time education, because they were unable to project when these possibly distant events will occur.

**Recommendation:** Though it is difficult to make a solid recommendation based on so few cases, we suggest that a "Don't know" response option be added to this question, as the final response column.

Q89 **How many of your brothers and sisters (including twin, step-or half-) left high school before graduating?**

Q90 **Do you babysit or take care of your own child, younger brothers or sisters, or other relatives?**

Q91 **On the average, how many hours per day are you responsible for their care?**

Q93 **In your family, who makes most of the decisions on each of the following topics?**

Q94 **How true are the following statements for you and your parents?**

First Follow-Up nonresponse to these items is high, with nonresponse rates ranging from 13.1 percent to 18.2 percent. This level of nonresponse probably is not due to problems with the questions, but arises from the length of the questionnaire, the serial position of the items, and the timed nature of the questionnaire administration. In the Second Follow-Up Field Test dropouts were given about 15 additional minutes to complete the questionnaire when necessary; nonresponse to these items is not high for the field test.

However, the Field Test Student Questionnaire nonresponse rate for question 91 is relatively high even when an adjustment is made for serial position. This nonresponse may be due to the use of the term "on the average", which respondents may have difficulty applying to the situation described in the question, especially when the time engaged may vary considerably across days.

**Recommendation:** Retain Q89, Q90, Q92, Q93, and Q94 in their present form. Reducing the length of the questionnaire for the main study and giving dropouts slightly more time to complete the questionnaire when necessary should keep nonresponse to these items low. Modify the wording of Q91, as suggested for the Student Questionnaire, to "About how many hours each day are you responsible for their care?"

Q95 **Lots of things happen in families that may affect young people. In the last 2 years, have any of the following happened to your family?**

Nonresponse to the subitems of this question was relatively high among the field test dropouts. Between three and five respondents did not answer each of the subitems. The number of items may have discouraged respondents; respondents may also have skipped items which they should have answered "No".

First Follow-Up nonresponse was also relatively high, at 12.5 percent. Nonresponse may have been an effect of the length of the questionnaire and serial position of the item.

**Recommendation:** We suggest, with qualification because of the small number of respondents, that several subitems be deleted from this question.

**Explanation Box Preceding Q99 and Q100**

Several of the cognitive interview respondents did not understand the explanation, particularly the term "fundamental freedoms of expression". Rather than help introduce the questions about religion, the explanation probably confuses respondents.

**Recommendation:** The wording of the explanation should be simplified, possibly to: "The following two questions deal with the basic right of self-expression. Your answers will help us interpret the results of the survey. We hope you will answer every question, but you may skip any question you do not wish to answer."

We may also wish to consider reordering Q99 and Q100. Because of the present order, respondents may think that by "religious person" in Q100 we mean someone who attends religious services regularly (the subject of Q99). Some respondents may report in Q100 that they do not consider themselves to be religious because they reported in Q99 that they seldom attend religious services.

### 5.3 Responses to OMB Queries - Dropout Questionnaire

The Office of Management and Budget (OMB), in correspondence relating to the clearance document, requested that we conduct analyses of the data collected by the three items listed below. Because of the small number of dropouts surveyed during the field test, it is impossible to make recommendations for the main study based on such analyses. Two of the three questions are also Student Questionnaire items and are the subject of analyses requested by OMB for that instrument; for these items we refer the reader to Section 3.8.

**Analysis: Q24** **How many times did the following things happen to you during the last semester you completed in school?**

This question was fielded in open-ended form, at the behest of the OMB reviewer, for the purpose of response category construction. The analogous Student Questionnaire item (Q8) was fielded in both open-ended and categorical response formats, for the same purpose.

**Recommendation:** Given the larger number of respondents completing the Student Questionnaire, these data will be used to inform Q24 of the Dropout Questionnaire. (See Section 3.4, Student Questionnaire Experiments, for a discussion of category construction.) We additionally recommend that this item be included if further cognitive interviewing or focus group testing of the Dropout Questionnaire is approved and undertaken during questionnaire revision before the main study.

**Analysis: Q29** **Please write below who referred you to this alternative program (such as a school counselor or principal).**

The OMB reviewer recommended that this question be fielded in open-ended form and that the data collected inform the construction of response options. Again, the small number of respondents limits the possibilities for empirical category construction. Of the nine dropouts who answered Q29, three reported that a school counselor had referred them to the alternative program; two by a principal or assistant principal; one by a friend; two by his or her parents; one by "the school"; and one respondent reported that he decided to enter the program himself. (Multiple responses were given to this question.)

**Analysis: Q45** **How much money do you think a person who works in the following professions earns per year?**

The OMB reviewer suggested that this question, corresponding to Q69 of the Student Questionnaire, be fielded in open-ended form for the purpose of constructing response categories.

**Recommendation:** We recommend that the categories constructed using the data from the parallel student item also be used for Q45 of the Dropout Questionnaire, (see Section 3.8, Response to OMB Queries - Student Questionnaire) and that the item be included in cognitive interview or focus group testing if such is approved.

### 5.4 Recommended Changes to the Dropout Questionnaire

**Item Refinement**: In Section 5.2.3 we discussed in detail the problematic items identified by analysis of cognitive interview findings, Second Follow-Up Field Test data, and First Follow-Up Dropout Questionnaire data. Four general recommendations emerged from this discussion: (1) clarify vague, difficult, or abstract language; (2) simplify questions and skip instructions when possible; (3) make question format readily intelligible; and (4) provide respondents with clear instructions for recording responses.

Throughout the Dropout Questionnaire text must be simplified. In its present form, the language used in some question stems, response options, instructions, and explanations appears to be inappropriate to the reading level of the respondents. Phrases such as "fundamental freedoms of expression" and "the relation of your training and education to your work experience" should be rewritten to make them more intelligible to dropouts. Wording that may appear to be precise, such as "your high school program", but which may be vulnerable to misinterpretation ("your high school's program"), should be rephrased to eliminate this possibility. When necessary and feasible, examples of terms which cannot be clarified further should be given.

Question content and format, along with skip instructions, must be simplified. Several items in the Dropout Questionnaire attempt to collect too much information in one question. Respondents become confused by the density of response options or the crowding of subitems, and either answer the question incorrectly or skip it. When such items act as filters, the data quality of the dependent questions suffers. We strongly recommend that these crowded or poorly-designed questions be reformatted, in some cases by dividing one question into two or more separate, more focused items.

Finally, dropouts must be given more guidance regarding correct responding behavior. This additional guidance could take the form of more explicit instructions within items, to reduce such respondent errors as multiple response selection for single response questions. Expansion of the general instructions may also be advisable, along with increased verbal instruction from team leaders.

(In addition to these four recommendations specific to the Dropout Questionnaire, we strongly recommend the conversion of all questions with a "mark all that apply" response format to a format requiring a "yes" or "no" response to each subitem, for reasons outlined in Section 3.4.4.)

**Questionnaire Length:** The reading skills of dropouts are generally poorer than those of their student peers; optimal questionnaire length is a critical issue.

Dropouts receive more attention from team leaders during group off-site survey administrations than students do because of the smaller group size compared to in-school sessions. Team leaders have more time to answer respondents' questions, to identify respondents having difficulty with items, and to retrieve critical items. Timing of the Dropout Questionnaire administration is also more relaxed, since the length of the survey period is not limited by school restrictions. However, dropouts cannot be given unlimited time to complete the questionnaire. Group off-site administrations should mirror Survey Day to the extent possible to ensure comparability of results and to limit costs. In the field test dropouts were given about fifteen minutes extra to finish the questionnaire when necessary.

For the small number of dropouts surveyed during the field test, nonresponse rates for questions at the end of the questionnaire do not appear to be significantly higher. However, from the dropout frequencies, response to terminal questions in the First Follow-Up instrument appears to have been adversely affected by item serial position. Consequently, we strongly recommend that the questionnaire be shortened slightly, from the current 100 questions to 90 or 95 items.

**Cognitive Interview Iterations:** We strongly recommend testing the revised questionnaire with additional cognitive interviews. Completion of another ten such interviews late in the questionnaire revision process would be extremely valuable and help to ensure that the changes made have solved the problems identified.

## Chapter 6: Parent Survey Procedures

This chapter describes the data collection and preparation procedures for the parent component of the Second Follow-Up Field Test. Parent Questionnaires were mailed to selected parents of student cohort members in mid-February 1991. Telephone follow-up with nonrespondents continued through the first week of April. Prior to the elaboration of mail-out and follow-up procedures, questionnaire and mail experiments implemented in the Parent Survey and selection of the field test parent sample are discussed.

### 6.1 Objectives, Content and Experiments in the Second Follow-Up Field Test Parent Questionnaire.

**Objectives of the Parent Questionnaire:** The central role played by the family and the home environment in shaping motivations and reinforcing learning -- and in molding habits, skills and aspirations -- makes family background information a key element in the analysis of NELS:88 data.

In an attempt to gather information on family factors that influence educational growth and failure, the Second Follow-Up parent questionnaire builds on and extends the range of topic areas covered by the HS&B parent questionnaire and the NELS:88 Base Year parent questionnaire.

The development of the parent survey instrument was guided by several principles, in addition to the methodological postulates of sound questionnaire design: specifically, these principles are:

- to cover major issues of policy related to the effects of families on the school achievement and educational attainment of students;
- to cover major issues of policy related to the effects of families on the educational and occupational outcomes of the out-of-high school population of dropouts and early graduates;
- to bridge the gap in data between Base Year and Second Follow-Up that resulted from lack of a parent survey in the First Follow-Up;
- to maximize comparability with items from the Base Year and earlier studies in order to facilitate within- and cross-cohort comparisons, longitudinal assessment, and analysis of trends; and
- to give due recognition to the new content domains that become relevant with the new transition point represented by completion of high school.

These considerations are elaborated upon below.

A key objective of the parent survey was to cover issues of relevance to the in-school population, as well as issues pertaining to parents of out-of-school sample members. Thus a single instrument is required to elicit information on topics ranging from the roles played by parents and families in the lives of school dropouts, at the one end of the spectrum; to, at the other end, the ways that students and their families make decisions regarding, and financially prepare for, the costs of pursuing postsecondary schooling.

**Content of the Parent Questionnaire:** For many families, planning for their child's postsecondary education or the transition directly to the world of work may already have begun in tenth grade (if not before). Many families also may have changed significantly in composition or situation between the 1988 Base Year and the 1992 Second Follow-Up. An additional burden, then, that must be assumed by the parent questionnaire is that of bridging a gap of fully four years.

For most parents there will at least be Base Year data. However, there are three distinct populations for whom there is no baseline parent data: Base Year parent nonrespondents, parents of First Follow-Up freshened students, and parents of Second Follow-Up freshened students. This fact requires that key static classification variables be systematically repeated on the Second Follow-Up parent questionnaire, or that such variables be specially repeated only for parents new to the study.

The parent questionnaire will build on the Base Year and HS&B instruments by repeating key intra- and cross-cohort comparison variables. In recognition of the changed policy agenda, and to accommodate the advance in age and stage of the student sample since 1988, it will also add a considerable body of new material. The fact that the financing of postsecondary education remains a crucial determinant of entry into and completion of postsecondary education serves as a potent reminder of the fact that the second semester of the senior year marks the study's entry into a second major transition point. Thus while the parent questionnaire will continue to examine the home educational support system, the support given to postsecondary plans and plans for entering the labor force will take their place as issues beside parental involvement in the high school and parental activities to monitor and support the child's learning.

Items included in the questionnaire for parents fall roughly into the following domains:

- family background (including racial, ethnic and socioeconomic characteristics);
- family attitudes towards education;
- non-English use patterns in the home;
- literacy supports and role models in the home;
- parent educational expectations for the child;
- parent involvement in, knowledge of, and satisfaction with the child's schooling and school environment;
- parent/child interactions;
- family decision-making;
- family planning for the transition to work; and
- financial and educational cost planning for postsecondary education.

Issues of inter-parent/school integration ("intergenerational closure") are also a topical focus of the Second Follow-Up parent survey.

### Parent Questionnaire and Mail Experiments

**Item order experiments:** Questionnaire experiments were conducted to assess the impact of item order on responses to two important parent reports -- satisfaction with their teenager's high school education and their teenager's alcohol and drug use. In one experiment, parents' reports of satisfaction with their teenager's high school education were made either before or after parents rated the school on several specific qualities. In a second experiment, parents' reports of their teenager's alcohol and drug use were made either before or after parents reported on the alcohol and drug use of their teenager's friends. The experiments were conducted to assess which ordering of items is likely to facilitate complete and unbiased responses. See section 7.3 for results of these questionnaire experiments.

**Mail experiments:** Two manipulations of materials in the questionnaire package were implemented in the parent component of the field test. These were: (1) alternative cover letters and (2) a direct reference to NELS:88 on the mailing envelope. Because mail surveys can suffer from low compliance rates, the purpose of these manipulations was to determine how to increase response rates to the mail-back portion of the parent survey.

Two versions of the cover letter were prepared and sent to parents. One version of the letter closely resembled the content and form of the Base Year Parent cover letter. The second letter covered the same content but used a crisp declarative sentence style, with bulleted items highlighting the importance of completing all of the questionnaire items and promptly mailing the document to NORC.

The field test also assessed the effectiveness of a direct reference to NELS:88 on the envelope used for mailing the questionnaire packets. See Section 6.3 for further details on Parent Survey data collection procedures. See Section 7.7 for the results of the mail experiment.

## 6.2 Sample Selection for the Parent Survey

To reduce the cost of the Parent Survey field test by collecting only as many completed questionnaires as were necessary for an assessment of the instrument and for analysis of the questionnaire experiments, a subsample of parents of 2FU field test students was chosen. This section describes how that subsample of parents was selected.

The initial pool of schools from which the parent sample was drawn included 70 field test schools that agreed to hold in-school survey sessions. A total of 24 schools were not included in this pool. Schools were excluded for two reasons: the school refused the parent component altogether (nine schools), or the school agreed to the parent survey but refused to provide NORC with addresses to mail questionnaires to parents (15 schools).

The goal in selecting the sample for the NELS:88 Second Follow-Up Parent Survey Field Test was to obtain a sample size of approximately 650 students from schools with a high degree of overlap with students whose parents were eligible for the Base Year Field Test parent sample. It was desirable to adequately represent these students in our sample because they reflect the majority of respondents in the main survey -- those who were participants in the Base Year survey and whose parents were eligible for inclusion in the Parent Survey.

The number of schools for the Second Follow-Up Field Test Parent Survey was determined by first calculating the average number of sampled students per school (assuming one parent would be selected for each student) and dividing the desired sample size by the average number of students per school. Schools were then stratified by control (public v. private). The selection probability for each school took into account the number of parents in the school who were eligible for the Base Year Field Test Parent Survey sample.

This procedure resulted in a sample of 18 schools and 688 parents. Of these, 244 (35 percent) were selected into the Base Year Field Test Parent Survey sample. Thirty-five (5 percent) parents had children who were in the Base Year field test sample but were not part of the Base Year field test parent survey sample. The remaining sample members were parents of students who were either freshened First or Second Follow-Up Field Test students or who were added to the 2FU field test sample to aid the psychometric assessment of the cognitive tests. Once parents were selected into the second follow-up field test parent sample they were randomly assigned to receive one of four versions of the questionnaire, and to be in one of four versions of the materials experiment.

### 6.3 Data Collection

**Mailout Procedures:** A Parent Questionnaire packet was mailed to each of the 688 parent survey sample members on February 12th. The packet consisted of a copy of one of the four versions of the Parent Questionnaire, a letter introducing the study to the parent, a pencil, and a postage-paid envelope for the return of the questionnaire. The letter explained the purpose of the study, with emphasis on the Parent Survey and its confidential nature, and requested that the parent promptly return the completed Parent Questionnaire to NORC. The letter instructed respondents to call NORC collect if they had any questions about the study or the questionnaire.

Packets mailed to Base Year Parent Survey nonparticipants (including parents of freshened and augmented students) were addressed "TO THE PARENTS OF: [STUDENT NAME]". Packets for Base Year participants were addressed to the Base Year parent respondent. Parent Questionnaires were labeled as just described, as well. The introductory letter explained that, if a particular person (the Base Year respondent) were named on the questionnaire label, that parent or guardian should complete the questionnaire, and that if the questionnaire were addressed "TO THE PARENTS OF...." the parent or guardian most knowledgeable about the student's current school situation and educational plans should fill out the questionnaire. The addresses used for the Parent Questionnaire mailout were collected by NORC representatives during their fall visits to schools.

Two experiments involving presentation of material were conducted as part of the mailout of questionnaires to parent subsample members. These experiments were designed to investigate methods for increasing parent participation rates in the survey.

For the first experiment, two versions of the letter accompanying the Parent Questionnaire were developed. One version was very similar to the Base Year parent survey cover letter -- a standard dense-text letter. The second version, identical to the first in content, used a different format -- a crisp declarative-sentence style, with bulleted items highlighting the importance of the parent survey and outlining the request.

For the second experiment, two versions of the mailing envelope were designed. One version was the standard brown NORC envelope, devoid of project identification. The second version was a white envelope with the NELS:88 logo printed in the lower left corner, and "NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988 Second Follow-Up" printed to the right of the logo.

With the two versions of the cover letter and two of the envelope, there were four letter-envelope combinations. Parent cases were randomly assigned to one of these four combinations, with the assignment recorded in the Survey Management System (SMS). (See Section 7.7 for the results of the mail experiment.)

Given the four Parent Questionnaire versions, there were sixteen different versions of the questionnaire packet, necessitating strict quality control. Questionnaire version and letter-envelope combination assignments were printed on the labels to be affixed to the questionnaire and envelope. To ensure that no parent was mailed the incorrect letter/envelope or questionnaire, the clerical staff were given the materials for only one of the sixteen packet versions at a time. Project staff checked the contents of each of the 688 packets prior to mailing, to ensure that parents received the correct versions of the materials.

**Prompting:** A total of 113 Parent Questionnaires were received at NORC by February 27th, 15 days after the packet mailout. The remaining 575 cases were prepared for nonresponse follow-up, which was conducted by telephone interviewers at NORC's Lake Park Data Collection Facility and by field interviewers who had served as team leaders during student data collection.

The primary purpose of nonresponse follow-up was to prompt parents to return completed Parent Questionnaires to NORC. Two prompts were made; the first generally occurred between March 4th and 11th, the second between March 12th and 19th. Interviewers were instructed to conduct telephone interviews during the prompting phase with respondents who refused to complete a self-administered questionnaire.

Interviewers attempted to conduct telephone interviews with respondents who failed to return completed questionnaires after the second prompt. The telephone interviewing phase of the follow-up effort occurred between March 20th and and April 8th.

The prompting manual provided to all interviewers included a general overview of NELS:88 and its follow-up studies, a section-by-section overview of the Parent Questionnaire, a case materials review, scripts for prompting calls, answers to anticipated parent questions, and refusal aversion/conversion suggestions. The Q-by-Q specifications guided the interviewer in converting the Parent Questionnaire -- designed for self-administration -- to a telephone-administered instrument. These specifications were developed to maximize comparability between self- and telephone-administration.

The two prompting telephone calls were designed to be a low-pressure reminder to the parent to complete and return the Parent Questionnaire. The initial prompting call also served as an opportunity to discover which parents needed to have another questionnaire packet mailed, to convince recalcitrant parents to participate in the study, and to trace initially unlocatable parents.

Interviewers used two prompting scripts for the initial prompting call, one for Base Year participants, the other for nonparticipants. For the former, interviewers were instructed to ask to speak to the Base Year parent respondent, determine whether he or she had yet completed and returned the questionnaire, and, if not, have the respondent give a specific date on which the questionnaire would be mailed. If the respondent reported that the questionnaire had already been mailed, the interviewer thanked the respondent for contributing to the study.

When the Base Year respondent was unavailable to complete the questionnaire within the follow-up period or refused to participate, the interviewer was instructed to prompt the other parent to complete the questionnaire, if he or she were knowledgeable about the student's current school situation and educational plans.

In the case of Base Year nonparticipants, interviewers were instructed to ask for the parent/guardian of the student most familiar with the child's current school situation and educational plans. After determining the identity of the respondent, the interviewer prompted him or her to complete and return the questionnaire.

Only minimal tracing of respondents was performed. Interviewers were instructed that the only locating resource they could use without supervisory approval was Directory Assistance. Once all locating problems had been identified at Lake Park and in the field, the Lake Park cases were batched by school and sent to the field interviewers, who contacted schools and attempted to update the case address information.

**Interviewing:** Given the brevity of the follow-up period, interviewers were instructed to conduct telephone interviews with respondents who had not returned completed questionnaires within one week of the second prompt. Field interviewers also conducted a limited number of personal interviewers during this period. Telephone interviewing occurred between March 20th and April 8th. (Interviewers also conducted a few telephone interviews during prompting, with respondents who refused to complete a self-administered questionnaire, but were receptive to a telephone interview.)

Interviewers administered a telephone-administration adaptation of the Parent Questionnaire. Interviewers were instructed to closely follow the instructions given in the question-by question specifications, in order to ensure comparability of data collected by telephone interviews and self-administered questionnaires. In addition, to approximate self-administration as much as possible, interviewers strongly recommended to respondents that they follow along on their copy of the questionnaire during the telephone interview, if possible.

Respondents who refused to participate in a telephone interview and insisted that they would return a self-administered questionnaire were asked to give the date on which they expected to mail the questionnaire. Respondents requesting remails during the interviewing period were promptly sent questionnaires.

Interviewers were instructed to ask respondents who said that they had already returned the questionnaire for the date that it was mailed. If the questionnaire had not been received at NORC within four to five days of that date, interviewers called the respondent again and attempted to conduct a telephone interview.

Field interviewers were instructed to schedule personal interviews only if they had confirmed that the respondent had no phone and were almost certain that a personal visit would result in a completed case. Travel time for personal interviews was limited to two hours roundtrip. Lake Park cases requiring personal interviewing were transferred to field interviewers, when necessary.

### 6.4 Data Preparation

Data preparation encompasses all activities associated with converting data from completed Parent Questionnaires to data files. These activities included receipt, manual edit, coding and data entry.

**Receipt Control:** Parent Questionnaires were receipted in the Receipt Control Shop of the NORC Lake Park Facility. Self-administered questionnaires were assigned the appropriate mode of administration disposition upon receipt; this disposition was not entered into the SMS, but was keyed during data entry. Telephone-administered questionnaires were assigned the appropriate mode disposition by the interviewer completing the case.

Receipt control clerks updated parent case dispositions in the Survey Management System (SMS) on a flow basis. Parent dispositions were again updated when questionnaires were sent to data entry. Each disposition transaction was automatically dated by the SMS.

**Editing/Coding:** A total of 506 parent questionnaires were edited for missing items in the Edit Shop of NORC's Lake Park Facility. Manual edits, including insertion of standard reserve codes, were performed as described in Section 2.3.

**Data Entry:** Two file layouts (essentially column position and field length for every item in the questionnaire) were developed for data entry of the questionnaires, one for forms A and C, the other for forms B and D. NORC's data entry subcontractor, BSI, prepared a test file of the data for review, keying five of each of the four forms.

The remainder of the Parent Questionnaires were then sent for data entry. Questionnaire data were converted to machine-readable form using a conventional key-to-disk method. Data were fully verified (that is, re-entered independently by a different operator), with the software comparing all fields for exact match. Because the primary purpose of the field test dataset is to help identify problems in the logic and content of the questionnaire, no machine editing (logical consistency cleaning) was performed.

## Chapter 7: Analysis of Parent Survey Data

This chapter discusses the results of an in-depth analysis of the Parent Questionnaire data. Parent response rates -- overall and by participation in previous waves -- are briefly examined. In subsequent sections, specific recommendations are offered for improving the effectiveness of the questionnaire in collecting high quality data, based on (a) item nonresponse analysis, (b) logical consistency analysis of responses to filter and dependent questions, and (c) analyses requested by the Office of Management and the Budget in correspondence relating to the clearance document.

The results of questionnaire experiments -- investigating item order effects -- are discussed and analyzed, along with the results of a procedural experiment involving questionnaire mailing materials. In the final section of the chapter, general recommendations for changes in the instrument and also in data collection procedures are presented. A modified data collection plan is described, the purpose of which is to facilitate the collection of high quality data concerning postsecondary education plans.

### 7.1 Parent Response Rates

This section looks at rates of parent participation in the NELS:88 Second Follow-Up (NELS:88 2FU) Field Test. Parent participation in the 2FU Field Test was defined as completion of the Parent Questionnaire. Response rates were derived by dividing the number of subsampled parents who completed a questionnaire by the number of parents with whom data collection was attempted.

Overall, 74 percent of all sample members with whom data collection was attempted completed a parent questionnaire. Of the 506 completed cases, 61 percent were self-administered and returned by mail, 38 percent were completed by telephone administration and 2 cases were completed as in-person interviews. Telephone interviewers encouraged respondents to read along on their copies of the questionnaire during the interview; however, this occured in less than 10 percent of the telephone interviews.

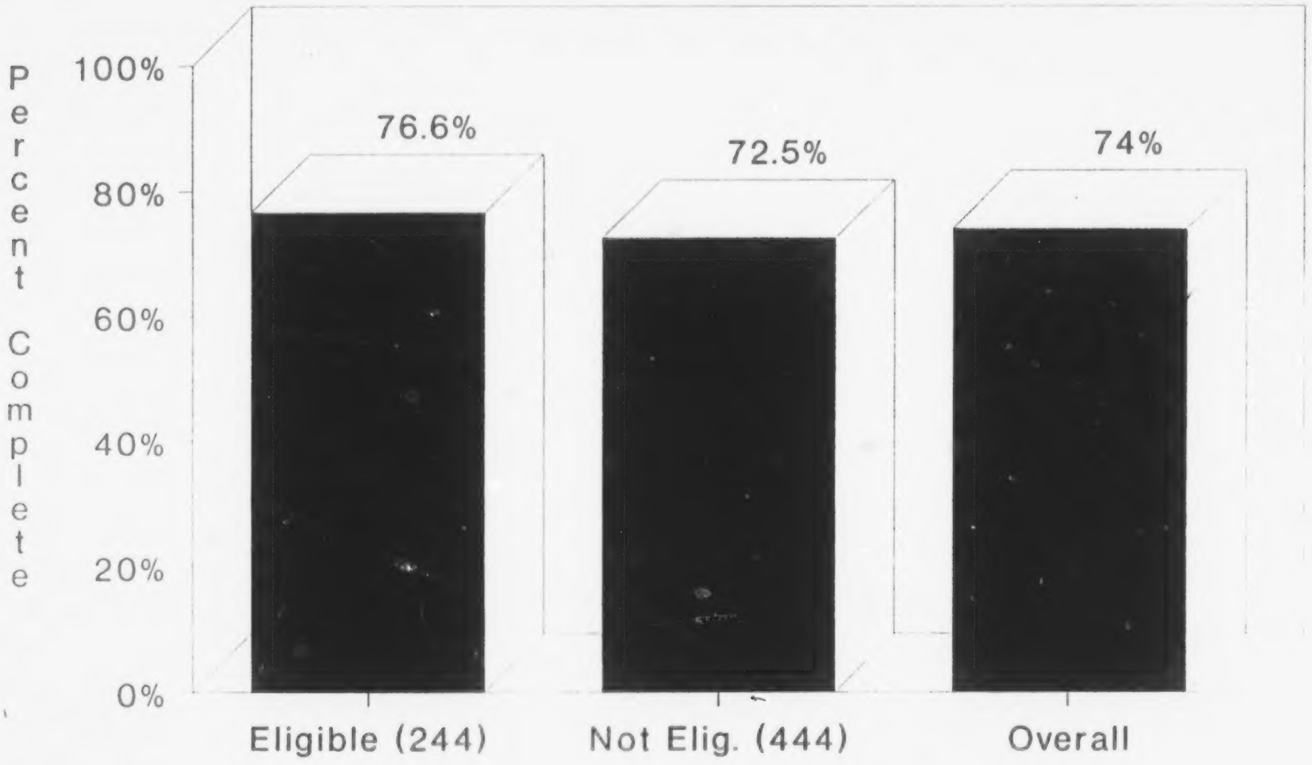
Figure 7.1 displays response rates by parents' participation in the Base Year Parent Survey. As figure 7.1 illustrates, parents who were eligible to participate in the Base Year Field Test (that is, parents of cohort members) were more likely to participate in the 2FU Field Test. Specifically, 77 percent of those eligible to participate in the Base Year study completed a 2FU parent quesionnaire. In comparison, only 72 percent of those who were not eligible to participate in the Base Year study (parents of freshened and augmented students) took part in the 2FU Field Test.

The fact that a higher response rate was attained for parents of cohort members is encouraging since a majority of the parent sample for the 2FU main study will be comprised of parents of core sample members. However, it will also be necessary to elicit high levels of response from other segments of the parent sample. These segments will include parents of dropouts and parents of freshened students. Convincing these parents of the importance of NELS:88 and gaining their cooperation will be a major challenge facing the 2FU main study parent survey.

### 7.2 Parent Item Nonresponse

In this section we present an analysis of item nonresponse in the Parent Questionnaire. Because of the high number of partial telephone interviews conducted and the expectation that the majority of Parent Questionnaires will be self-administered in the main study, only respondents who completed self-administered questionnaires (a total of 311) were included in this item nonresponse analysis. In

# Figure 7.1.1. - Parent Completion Rates by Base Year Eligibility



n = 688
Source: NELS:88 2FU Field Test Data

addition to identifying items with high nonresponse rates, we present recommended wording and format changes likely to improve response rates.

**Overview of the results.** Few questions in the Parent Questionnaire exhibit high rates of nonresponse. For those that do, we have identified four sources of item nonresponse: (1) nonresponse due to poor question design, (2) nonresponse arising from overcrowded questions, (3) nonresponse due to difficult or unclear language, and (4) nonresponse due to conditioned response behavior that is inappropriate for the response format of the item.

One example of the first source is Q70, in which a skip item is embedded as the first response in the question matrix, disrupting the continuity between the question stem and the remaining response categories.

A striking example of the second source of nonresponse, overcrowded questions, can be found in Q54, which inquires about parents' encouraging their teenagers to prepare for the SAT, ACT, or ASVAB. The question format is unconventional, with subitem labels heading the columns, and the question is so crowded that it is difficult for respondents to follow the logic of the format. We suggest that for this and similar questions the format be simplified and/or the question be divided into two or more questions. Question 25 is a good example of the third source of nonresponse. The question stem is unnecessarily long, and respondents lose their way among the clauses.

Examples of the last source can be found in Q25, Q26, Q70, Q75, and Q79. The use of a "mark all that apply" response format in prior items in the Parent Questionnaire may condition respondents to answer these questions, which require a "yes" or "no" response to each subitem, as though they had a "mark all that apply" format. We recommend that all items with a "mark all that apply" response format be revised to a "Yes/No" format, since the latter format elicits higher quality data (See section 3.4.4).

**Calculation of item nonresponse rates.** Parent Questionnaire item nonresponse rates were calculated in three steps. First, dichotomous variables were created for each questionnaire item, with "1" indicating no response was given for the item and "0" indicating a valid response was given. Then, respondents who legitimately skipped dependent questions were coded as missing for those questions. Finally, item nonresponse rates were derived by calculating the mean of the dichotomous variables. The resulting item nonresponse rates obtained by computing the mean of the dichotomous variables, indicate the proportion of respondents eligible to answer an item who failed to answer the item. As noted above, nonresponse rates were calculated using only the 311 respondents who completed self-administered questionnaires. Item nonresponse rates greater than ten percent were considered to be high. Item nonresponse rates for items meeting this criteria in the Parent Questionnaire are displayed in Table 7.2.1. Ten percent rather than 20 percent (used in the student analysis) was used as the nonresponse threshold because the Parent Questionnaire was not timed and respondents could complete it at their leisure.

Since the Parent Questionnaires were self-administered, we made no adjustment in calculating nonresponse rates to control for the effect of questionnaire length as with the Student Questionnaire. All response rates reported are unadjusted, unless noted.

Analysis: **Q10A What kind of job did or does he or she have?**
**Q10B What kind of business or industry was this job in?**
**Q10C What were/are his/her main activities or duties on this job?**
**Q10D In his or her present or most recent job is/was your spouse/partner:**

## TABLE 7.2.1

### Parent Questionnaire Item Nonresponse
### Items with High Nonresponse Rates

| Item | Multiple Response | Refusal | Don't Know | Missing | Nonresponse Rate |
|---|---|---|---|---|---|
| Q10A* | N/A | 0.00 | 0.00 | 0.12 | 0.12 |
| Q10B | N/A | 0.00 | 0.00 | 0.14 | 0.14 |
| Q10C | N/A | 0.00 | 0.00 | 0.18 | 0.18 |
| Q10D | 0.00 | 0.00 | 0.00 | 0.14 | 0.14 |
| Q13A | 0.00 | 0.00 | 0.00 | 0.23 | 0.23 |
| Q24OV | 0.00 | 0.00 | 0.00 | 0.40 | 0.40 |
| Q25A | 0.00 | 0.00 | 0.00 | 0.56 | 0.56 |
| Q25B | 0.00 | 0.00 | 0.00 | 0.63 | 0.63 |
| Q25C | 0.00 | 0.00 | 0.00 | 0.63 | 0.63 |
| Q25D | 0.00 | 0.00 | 0.00 | 0.69 | 0.69 |
| Q25E | 0.00 | 0.00 | 0.00 | 0.56 | 0.56 |
| Q25F | 0.00 | 0.00 | 0.00 | 0.63 | 0.63 |
| Q25G | 0.00 | 0.00 | 0.00 | 0.69 | 0.69 |
| Q25H | 0.00 | 0.00 | 0.00 | 0.63 | 0.63 |
| Q25I | 0.00 | 0.00 | 0.00 | 0.69 | 0.69 |
| Q25J | 0.00 | 0.00 | 0.00 | 0.69 | 0.69 |
| Q25K | 0.00 | 0.00 | 0.00 | 0.75 | 0.75 |
| Q25L | 0.00 | 0.00 | 0.00 | 0.69 | 0.69 |
| Q25M | 0.00 | 0.00 | 0.00 | 0.56 | 0.56 |
| Q25N | 0.00 | 0.00 | 0.00 | 0.56 | 0.56 |
| Q25O | 0.00 | 0.00 | 0.00 | 0.69 | 0.69 |
| Q26A | 0.00 | 0.00 | 0.00 | 0.47 | 0.47 |
| Q26B | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26C | 0.00 | 0.00 | 0.00 | 0.44 | 0.44 |
| Q26D | 0.00 | 0.00 | 0.00 | 0.47 | 0.47 |
| Q26E | 0.00 | 0.00 | 0.00 | 0.47 | 0.47 |
| Q26F | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26G | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26H | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26I | 0.00 | 0.00 | 0.00 | 0.47 | 0.47 |
| Q26J | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26K | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26L | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26M | 0.00 | 0.00 | 0.00 | 0.50 | 0.50 |
| Q26N | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26O | 0.00 | 0.00 | 0.00 | 0.53 | 0.53 |
| Q26P | 0.00 | 0.00 | 0.00 | 0.47 | 0.47 |
| Q36I | 0.00 | 0.00 | 0.00 | 0.13 | 0.13 |

* Denotes a critical item.

## TABLE 7.2.1 - continued

### Parent Questionnaire Item Nonresponse
### Items with High Nonresponse Rates

| Item | Multiple Response | Refusal | Don't Know | Missing | Nonresponse Rate |
|---|---|---|---|---|---|
| Q54SAT | 0.00 | 0.00 | 0.00 | 0.11 | 0.11 |
| Q54ACT | 0.00 | 0.00 | 0.00 | 0.19 | 0.19 |
| Q54ASVAB | 0.00 | 0.00 | 0.00 | 0.34 | 0.34 |
| Q68A | 0.00 | 0.00 | 0.00 | 0.13 | 0.13 |
| Q68B | 0.00 | 0.00 | 0.00 | 0.15 | 0.15 |
| Q68C | 0.00 | 0.00 | 0.00 | 0.31 | 0.31 |
| Q68D | 0.00 | 0.00 | 0.00 | 0.64 | 0.64 |
| Q70A* | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q70B* | 0.00 | 0.00 | 0.00 | 0.29 | 0.29 |
| Q70C* | 0.00 | 0.00 | 0.00 | 0.30 | 0.30 |
| Q70D* | 0.00 | 0.00 | 0.00 | 0.26 | 0.26 |
| Q70E* | 0.00 | 0.00 | 0.00 | 0.31 | 0.31 |
| Q70F* | 0.00 | 0.00 | 0.00 | 0.26 | 0.26 |
| Q70G* | 0.00 | 0.00 | 0.00 | 0.25 | 0.25 |
| Q70H* | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q70I* | 0.00 | 0.00 | 0.00 | 0.28 | 0.28 |
| Q70J* | 0.00 | 0.00 | 0.00 | 0.30 | 0.30 |
| Q75A | 0.00 | 0.00 | 0.00 | 0.14 | 0.14 |
| Q75B | 0.00 | 0.00 | 0.00 | 0.12 | 0.12 |
| Q75C | 0.00 | 0.00 | 0.00 | 0.14 | 0.14 |
| Q75D | 0.00 | 0.00 | 0.00 | 0.13 | 0.13 |
| Q75E | 0.00 | 0.00 | 0.00 | 0.07 | 0.07 |
| Q75F | 0.00 | 0.00 | 0.00 | 0.16 | 0.16 |
| Q75G | 0.00 | 0.00 | 0.00 | 0.14 | 0.14 |
| Q75H | 0.00 | 0.00 | 0.00 | 0.64 | 0.64 |
| Q79A | 0.00 | 0.00 | 0.00 | 0.17 | 0.17 |
| Q79B | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q79C | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q79D | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q79E | 0.00 | 0.00 | 0.00 | 0.21 | 0.21 |
| Q79F | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q79G | 0.00 | 0.00 | 0.00 | 0.24 | 0.24 |
| Q79H | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q79I | 0.00 | 0.00 | 0.00 | 0.23 | 0.23 |
| Q79J | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |

* Denotes a critical item.
Source: NELS:88 Second Follow-Up Field Test Data

This series of questions collects information about the job of the respondent's spouse or partner. Nonresponse to these items, ranging from 12 percent to 18 percent, is high relative to rates found for immediately preceding and following items. The majority of this nonresponse arises from respondents who did not answer Q9, the filter question for Q10A through Q10D. These respondents apparently skipped Q9 because they have no spouse or partner (according to response in Q6), and appear from the unadjusted data to have skipped the dependent questions illegitimately. When these cases are recoded as missing for items Q10A through Q10D, nonresponse rates are low, ranging from four percent to 9 percent.

**Recommendation:** We suggest that a "Does Not Apply" response option be added to the filter Q9, to provide better routing instructions for respondents with no spouse or partner and to permit their exclusion when calculating item nonresponse rates.

Analysis: **Q13A Which of these best describes your background? (Asian or Pacific Islander)**

This question seeks to determine the nationality of Asian or Pacific Islander respondents. The nonresponse rate for Q13A is relatively high (23 percent). However, all 21 respondents who responded "Asian or Pacific Islander" to Q12 (the filter) answered this question. The five respondents who failed to answer the filter question and also did not answer Q13A account for the nonresponse to the question.

**Recommendation:** Retain the item as is.

Analysis: **Q24 For which of the following reasons did your teenager stop attending school for 21 or more consecutive school days?**

**Q25 Did you or your spouse/partner do any of the following when, in the last 2 school years, your teenager stopped attending school for 21 or more consecutive school days for a reason other than illness or vacation?**

**Q26 Did the school do any of the following when your teenager stopped attending school for 21 or more consecutive school days?**

Question 23 is the filter for Q24 through Q26, which seek to determine whether the student dropped out of school in the previous two school years, why he or she dropped out, and what courses of action the parents and school took in reaction to the student's dropping out. Nonresponse to the dependent questions is very high, above 50 percent for almost all of the subitems. However, the base for calculating nonresponse rates is quite small, consisting of only 15 to 16 respondents. Closer examination of the data reveals that the eight respondents who did not answer the filter question also did not answer most (if not all) of the dependent questions, inflating the nonresponse rates. Exclusion of these eight respondents yields lower, but still high, nonresponse rates for Q24 through Q26 and their subitems; rates range from 18 percent to 55 percent.

**Recommendation:** Though it is difficult to draw conclusions regarding sources of nonresponse to these items because of the very small number of eligible respondents, scrutiny of the questions suggests a number of ways to improve response. Question 23, the filter question, does not contain explicit routing instructions for parents responding "Yes". Inclusion of such instructions ("Please answer questions 24 through 26") might be helpful, given the intimidating length of the dependent questions.

The stems of the dependent questions also could be simplified. Question 24 might be revised to read, "Why did your teenager stop attending school?" The addition of a "Don't know" response option to this question would also be appropriate, since some parents may not know why their teenager dropped out of school. Question 25 could be simplified to "When your teenager stopped attending school, did you or your spouse/partner...."

In addition, respondents may have difficulty making the transition from a "mark all that apply" format in Q24 to the "Yes/No" formats of Q25 and Q26. Because of the lengthy list of subitems in each of the latter, respondents may behave as though these questions have a "mark all that apply" format, rather than responding to each subitem. Since higher data quality is associated with the "Yes/No" response format, as discussed in Section 3.4.4, Q24 should be reformatted to require a response to each subitem.

Finally, reducing the number of subitems in each question would improve item nonresponse. Some respondents are probably intimidated by the lengthy lists of response categories (17 in Q26) and either skip entire questions or answer only some subitems. The small number of parents answering Q24 through Q26 makes item elimination based on empirical data impossible.

Analysis: **Q36 In your family, who makes most of the decisions about each of the following topics?**

**Q36I If privileges should be taken away because my teenager has used alcohol or drugs.**

The nonresponse rate to subitem Q36I is high (13 percent) relative to the other subitems of Q36. The higher nonresponse may arise from the sensitive nature of the item and also from its inapplicability to respondents. Some parents may not have answered this item because the issue of whether privileges should be revoked because of drug or alcohol use has never arisen.

**Recommendation:** Retain item as is. The level of nonresponse does not warrant the introduction of a "Not Applicable" response column.

Analysis: **Q54 Have you ever encouraged your teenager to get a book, a manual, or a computer program, or to take a course that would help him/her to prepare for any of the following tests?**

Each of the subitems of this question, which correspond to the tests (SAT, ACT, and ASVAB), has a high nonresponse rate relative to preceding and following questions. The ASVAB item demonstrates the highest nonresponse (34 percent), the ACT the second highest at 19 percent, and the SAT the lowest (nonresponse rate of 11 percent). This nonresponse gradient is probably due to the ordering of the subitems and respondents' relative familiarity with the tests. The SAT is the first item and the most familiar of the three tests, the ACT is the second item and probably second in familiarity, and the ASVAB is the third item and likely the most unfamiliar.

The high overall nonresponse to the subitems of this question is probably due to its confusing, crowded, and unconventional format; the question attempts to do too much in one matrix. Rather than heading the rows, as is the convention, the subitem labels (SAT, ACT, and ASVAB) head the columns. The response category labels head the rows. In addition, the response options are too dense as the question is presently formatted.

**Recommendation:** A more intelligible, though more lengthy, format would be:

54. Have you ever encouraged your teenager to get a book or a computer program, or to take a course, that would help him/her to prepare for any of the following tests?

A. Scholastic Aptitude Test (SAT)

(CIRCLE ONE)

No . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 GO TO A1

Yes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2 SKIP TO B

A1. Why have you not encouraged your teenager to get a book or a computer program, or to take a course, to prepare for the SAT?

I have never heard of this test . . . . . . . . . . 1

I don't think he/she
needs to take this test . . . . . . . . . . . . . . . . 2

He/she needs to take this test,
but doesn't need to prepare for it . . . . . . . 3

I have another reason . . . . . . . . . . . . . . . . 4

B. American College Test (ACT)

(CIRCLE ONE)

No . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 GO TO B1

Yes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2 SKIP TO C

B1. Why have you not encouraged your teenager to get a book or a computer program, or to take a course, to prepare for the ACT?

I have never heard of this test . . . . . . . . . . 1

I don't think he/she
needs to take this test . . . . . . . . . . . . . . . . 2

He/she needs to take this test,
but doesn't need to prepare for it . . . . . . . 3

I have another reason . . . . . . . . . . . . . . . . 4

BEST COPY AVAILABLE

C. Armed Services Vocational Aptitude Battery (ASVAB)

(CIRCLE ONE)

No . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 GO TO C1

Yes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2 SKIP TO Q. 55

C1. Why have you not encouraged your teenager to get a book or a computer program, or to take a course, to prepare for the ASVAB?

I have never heard of this test . . . . . . . . . . 1

I don't think he/she needs to take this test . . . . . . . . . . . . . . . . 2

He/she needs to take this test, but doesn't need to prepare for it . . . . . . . 3

I have another reason . . . . . . . . . . . . . . . . 4

**Analysis: Q68 Do you currently have any of the following educational expenses for any of your children?**

**Q68A Private elementary or high school tuition and associated expenses**
**Q68B College tuition and associated expenses**
**Q68C Tutoring**
**Q68D Other**

All four subitems of this question have high nonresponse rates: Q68A, 13 percent; Q68B, 15 percent; Q68C, 31 percent; and Q68D, 64 percent. Nonresponders to these items probably were respondents who should have circled "No", but instead skipped the items; respondents may have thought that this response behavior was acceptable, given the precedent set by earlier "mark all that apply" questions.

**Recommendation:** Converting questions having a "mark all that apply" format to one requiring a" yes" or "no" response to each item should result in higher response to the subitems of question 68. Response to Q68C also might be improved by making it the second item in the subitem list; its juxtaposition with Q68A ("Private elementary or high school tuition and associated expenses") seems more appropriate than its present position following "College tuition and associated expenses".

**Analysis: Q70 Which of the following have you or your spouse/partner done to financially prepare for your teenager's education after high school?**

Nonresponse rates for items Q70A through Q70J, ranging from 20 percent to 31 percent, are much higher than those for surrounding items. At least part of this nonresponse is probably due to poor question design. Continuity between the question stem and the subitems is disrupted by the intervening skip instruction to respondents whose teenagers do not plan to attend postsecondary school. Also, in answering the subitems, each of which requires a "yes" or "no" response, respondents may behave as though they are answering a "mark all that apply" question and skip inapplicable items, because of the precedent set by earlier questions with such a response format.

**Recommendation:** We suggest that the intrusive skip instructions be removed from the question and be reformatted as a filter question immediately preceding question 70. The filter question would then read:

Does your teenager plan to continue his or her education?

No ....................................... 1 SKIP TO QUESTION 83

Yes ..................................... 2 GO TO QUESTION 70

Analysis: **Q75 Have you or your spouse/partner done any of the following to learn about applying for financial aid for further education for your teenager?**

Nonresponse rates for items Q75A through Q75G are relatively high, ranging from 7 percent to 16 percent. The nonresponse rate for the final subitem, Q75H--"Other (WRITE IN BELOW)"--is even higher at 64 percent. Once again, some respondents probably answered this question as though it had a "mark all that apply" format, circling "Yes" for the items that were applicable and skipping those that were not. The number of subitems and the position of the question at the end of the instrument also may have contributed to nonresponse.

**Recommendation:** The conversion of questions having a "mark all that apply" format to one requiring a "yes" or "no" response to each item will result in higher response to Q75.

Analysis: **Q79 For which of the following reasons has your teenager not applied for financial aid?**

Nonresponse rates for items Q79A through Q79J are high--from 17 percent to 22 percent. As in Q75, some respondents probably answered this question as though it had a "mark all that apply" format, circling "Yes" for the items that were applicable and skipping those that were not. Once again, the number of subitems and the position of the question at the end of the instrument also may have contributed to nonresponse.

**Recommendation:** The conversion of questions having a "mark all that apply" format to one requiring a yes or no response to each item will result in higher response to Q79.

## 7.3 Results of Parent Questionnaire Experiments

In this section we discuss the results of two Parent Questionnaire experiments. The experiments were conducted to assess the effect of item order (i.e., context) on responses to parents' reports of satisfaction with their teenager's high school education and to parents' reports of drug use and drinking behavior exhibited by their teenager and their teenager's friends.

### 7.3.1 Effect of Item Order on Parental School Satisfaction Reports

The first study examines the effect of item order on parents' reports of satisfaction with their teenager's high school education. Parents were assigned at random to two versions of the questionnaire, one in which Q27 (general satisfaction with teenager's high school education) preceded Q28a to Q28m (evaluations of specific aspects of the school) and one in which Q27 followed Q28a to Q28m.

Frequencies for Q27 under the two item order positions are shown in Table 7.3.1. The overall pattern of responses across the two positions is not significant. Ignoring the missing values, 83.4 percent of respondents give a somewhat or very satisfied response to the education satisfaction question when it is preceded by the 12 specific evaluations (11 of which are positively worded). In comparison, 77.5 percent of respondents gave a somewhat or very satisfied response to the education satisfaction when this question preceded the 12 specific evaluations.

**Table 7.3.1**

**Crosstabulation of General Satisfaction with Teen's High School Education by Item Order**

```
Item          Count  |General |Specific|
Order-->    Col Pct  |First   | First  |  Row
                     |    1.00|    2.00| Total
 Q27         --------+--------+--------+
Very Unsatisfied     |    24  |    19  |    43
                     |   9.7  |   7.4  |   8.5
                     +--------+--------+
Somewhat Unsatis.    |    21  |    17  |    38
                     |   8.5  |   6.6  |   7.5
                     +--------+--------+
Somewhat Satisfied   |    79  |    83  |   162
                     |  31.9  |  32.2  |  32.0
                     +--------+--------+
Very Satisfied       |    76  |    98  |   174
                     |  30.6  |  38.0  |  34.4
                     +--------+--------+
Missing              |    48  |    41  |    89
                     |  19.4  |  15.9  |  17.6
                     +--------+--------+
             Column      248      258     506
              Total     49.0     51.0   100.0

 Chi-Square     D.F.      Significance        Min E.F.      Cells with E.F.< 5
 ----------     ----      ------------        --------      ------------------

   4.23741        4          .3748             18.625         None
```

We treated the four-point satisfaction scale as a quasi-continuous variable and used Analysis of Variance (ANOVA) to test whether the difference was significant. ANOVA was selected in order to assess the effect of both item order and mode of administration (self-administration versus telephone), as well as of the interaction between item order and mode. Results are shown in Table 7.3.2.

## Table 7.3.2

## Effects of Item Order and Mode on Parents' Reports of Satisfaction with their Teenager's High School Education

TOTAL POPULATION

```
    3.12
(   414)


Item Order

  General   Specific
   First     First
    3.04      3.19
(   200) (    214)

Mode of Administration

    Self     Telephone
    3.11      3.13
(   329) (     85)

          MODE
              Self      Telephone
Item Order
 General      3.04       3.02
 First    (   158)  (     42)

Specific      3.18       3.23
First     (   171)  (     43)


                                   Sum of                 Mean
Signif
Source of Variation               Squares     DF        Square       F      of
F

Main Effects                         2.560      2         1.280    1.404
.247
  Item Order                         2.540      1         2.540    2.786
.096
  Mode                                .025      1          .025     .027
.869

2-way Interactions                    .072      1          .072     .079
.778
  Order   By Mode                     .072      1          .072     .079
.778

Explained                            2.632      3          .877     .962
.410

Residual                           373.803    410         .912

Total                              376.435    413         .911
```

The results suggest a marginal effect of item order. Placing the specific items before the general evaluation seems to induce some parents to respond more positively to the general evaluation item.

**Recommendation.** This result is consistent with the context effects literature demonstrating that prior items can affect responses to subsequent ones, especially when the subsequent item requires a general evaluation. We recommend that the general satisfaction item be placed before any items asking parents to evaluate specific aspects of the school.

### 7.3.3 Effect of Context on Parents' Reporting of Drug/Alcohol Problems.

Reporting the drug or alcohol problems of one's child is a sensitive issue. Nonetheless, it is useful for researchers to obtain parents' perspectives on this issue. Therefore, in the Parent Questionnaire several items have been included which ask parents to report on their child's alcohol and drug use (Q48a and Q48b). In addition, parents are asked to report on the drug and alcohol use and other delinquent behavior of their childrens' friends (Q49A to Q49E). It is believed that responses to drug and alcohol use among the teenagers' friends might serve as an indirect method for garnering from the parent whether such problems exist with the teenager.

When one considers the two sets of reports together, interesting possibilities for facilitating responding arise. Parents may be reluctant to admit to their own child's alcohol or drug use partly because parents may perceive that such a problem, if admitted, reflects badly on them. Putting these items in a larger social context, by asking parents first about their teenager's friends' alcohol and drug use, may reduce that threat and result in parents' increased willingness to report their own child's problems. Thus, asking parents about their child's friends' problems first may result in more drug and alcohol-related problem reporting in reference to their teenagers. On the other hand, parents may become so defensive when they are asked about their own child's problems in these areas that they become less willing to admit that their childrens' friends may have problems in these areas. Thus, asking parents about their child's problems first may result in less drug and alcohol-related problem reporting in reference to their teenagers' friends.

To test which of the orders of items facilitated eliciting reports about drug and alcohol problems from parents, two versions of the Parent Questionnaire were created. In one version parents received the two questions about their child's drug and alcohol use first. In the second version, parents first received the questions about their child's friends' drug and alcohol use, and other delinquent behavior.

Table 7.3.3. shows the frequencies of responses to the questions about the teenager's drinking and drug problems. As can readily be seen, parents are loathe to admit their teenagers have problems in these areas.

In order to assess the effect of item order on reports of teen drinking and drug use, Q48 and Q49 were collapsed into a binary variable, with one category representing the "no problem" response and a second category representing the positive responses and "don't know" responses. Missing values were excluded from the analysis. An item order (report on teen first versus report on teen's friends first) by mode (self-administered versus telephone) logit analysis was conducted. As might be expected from the extremely skewed distribution and small variation in responses, neither factor influenced responses, nor did the interaction influence responses.

Table 7.3.4. shows the frequencies of responses to the questions about drinking and drug problems and delinquency in reference to the teenager's friends. Parents are somewhat more willing to admit problems in reference to their teenager's friends than their teenager.

## Table 7.3.3
## Responses to questions about teen drinking and drug problems

Q48a. My teenager has a drinking problem

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| No Problem | 1.00 | 395 | 78.1 | 78.1 | 78.1 |
| Small Problem | 2.00 | 9 | 1.8 | 1.8 | 79.8 |
| Moderate Problem | 3.00 | 1 | .2 | .2 | 80.0 |
| Large Problem | 4.00 | 2 | .4 | .4 | 80.4 |
| Don't Know | 5.00 | 2 | .4 | .4 | 80.8 |
| Missing | 9.00 | 97 | 19.2 | 19.2 | 100.0 |
|  | TOTAL | 506 | 100.0 | 100.0 |  |

Q48b. My teenager has a drug problem

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| No Problem | 1.00 | 401 | 79.2 | 79.2 | 79.2 |
| Small Problem | 2.00 | 4 | .8 | .8 | 80.0 |
| Moderate Problem | 3.00 | 0 | .0 | .0 | 80.0 |
| Large Problem | 4.00 | 1 | .2 | .2 | 80.2 |
| Don't Know | 5.00 | 2 | .4 | .4 | 80.6 |
| Missing | 9.00 | 98 | 19.4 | 19.4 | 100.0 |
|  | TOTAL | 506 | 100.0 | 100.0 |  |

## Table 7.3.4
## Responses to questions about teen's friends' drinking, drug problems and other delinquent behavior

Q49a. Drinking is a problem among my teenger's friends

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| Strongly Disagree | 1.00 | 211 | 41.7 | 41.7 | 41.7 |
| Disagree | 2.00 | 130 | 25.7 | 25.7 | 67.4 |
| Agree | 3.00 | 38 | 7.5 | 7.5 | 74.9 |
| Strongly Agree | 4.00 | 8 | 1.6 | 1.6 | 76.5 |
| Don't Know | 5.00 | 22 | 4.3 | 4.3 | 80.8 |
| Missing | 9.00 | 97 | 19.2 | 19.2 | 100.0 |
| | TOTAL | 506 | 100.0 | 100.0 | |

Q49b. Drug use is a problem among my teenager's friends

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| Strongly Disagree | 1.00 | 250 | 49.4 | 49.4 | 49.4 |
| Disagree | 2.00 | 125 | 24.7 | 24.7 | 74.1 |
| Agree | 3.00 | 9 | 1.8 | 1.8 | 75.9 |
| Strongly Agree | 4.00 | 4 | .8 | .8 | 76.7 |
| Don't Know | 5.00 | 21 | 4.2 | 4.2 | 80.8 |
| Missing | 9.00 | 97 | 19.2 | 19.2 | 100.0 |
| | TOTAL | 506 | 100.0 | 100.0 | |

Q49c. Drinking among my teenager's friends has been a bad influence on my teenager

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| Strongly Disagree | 1.00 | 237 | 46.8 | 46.8 | 46.8 |
| Disagree | 2.00 | 130 | 25.7 | 25.7 | 72.5 |
| Agree | 3.00 | 21 | 4.2 | 4.2 | 76.7 |
| Strongly Agree | 4.00 | 7 | 1.4 | 1.4 | 78.1 |
| Don't Know | 5.00 | 14 | 2.8 | 2.8 | 80.8 |
| Missing | 9.00 | 97 | 19.2 | 19.2 | 100.0 |
| | TOTAL | 506 | 100.0 | 100.0 | |

Q49d. Drug use among my teenager's friends has been a bad influence on my teenager

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| Strongly Disagree | 1.00 | 270 | 53.4 | 53.4 | 53.4 |
| Disagree | 2.00 | 109 | 21.5 | 21.5 | 74.9 |
| Agree | 3.00 | 10 | 2.0 | 2.0 | 76.9 |
| Strongly Agree | 4.00 | 5 | 1.0 | 1.0 | 77.9 |
| Don't Know | 5.00 | 15 | 3.0 | 3.0 | 80.8 |
| Missing | 9.00 | 97 | 19.2 | 19.2 | 100.0 |
| | TOTAL | 506 | 100.0 | 100.0 | |

**Table 7.3.4 (cont.)**
**Responses to questions about teen's friends' drinking,**
**drug problems and other delinquent behavior**

Q49e. Theft and violence are a problem among my teenager's friends

| Value Label | Value | Frequency | Percent | Percent | Percent |
|---|---|---|---|---|---|
| Strongly Disagree | 1.00 | 273 | 54.0 | 54.0 | 54.0 |
| Disagree | 2.00 | 107 | 21.1 | 21.1 | 75.1 |
| Agree | 3.00 | 10 | 2.0 | 2.0 | 77.1 |
| Strongly Agree | 4.00 | 3 | .6 | .6 | 77.7 |
| Don't Know | 5.00 | 15 | 3.0 | 3.0 | 80.6 |
| Missing | 9.00 | 98 | 19.4 | 19.4 | 100.0 |
|  | TOTAL | 506 | 100.0 | 100.0 |  |

Analyses of variance were used to assess the effect of item order on parents' reports of the problems of their teenager's friends. This linear technique was chosen because of the greater variation in responses to these items than to the items about the teenager. Before submitting these variables to analysis, missing responses were excluded and "don't know" responses were recorded to a middle category. This had the result of maintaining a five point scale with "don't know" represented by "3" and the "1" and "5" endpoints represented by "strongly disagree" and "strongly agree" respectively.

First, a multivariate analysis of variance was conducted on five questions considered simultaneously. Significant effects for order (Multivariate $\underline{F}(5,395)=3.32$, $p<.001$) and mode (Multivariate $\underline{F}(5,395)=2.31$, $p<.05$) were found. The interaction between order and mode was marginally significant (Multivariate $\underline{F}(5,395)=2.02$, $p<.08$).

Next, separate univariate analyses of variance were conducted for Q49A to Q49E. A summary of the results is shown in Table 7.4.5. The interaction between order and mode is marginally significant for all but one item (Q49B) but the pattern of means in the order by mode cells is the same for all items. In general, parents are less willing to admit that their teenager's friends have problems when the two items about their teenager's problems precede the items about the friends' problems. This is especially true for those parents interviewed by telephone. This defensiveness may suggest that the items about the teenager's friends are indeed operating as indirect measures of the teenager's problems.

Table 7.4.5. Effects of Order and Mode on Parents' Reports of Teen's Friends' Delinquent Behavior

|  | MAIN EFFECTS |  |  |  | INTERACTION |  |  |  |
|---|---|---|---|---|---|---|---|---|
|  | ORDER |  | MODE |  | Self-Admin. |  | Telephone |  |
|  | Teen First | Friend First | Self-Admin. | Telephone | Teen First | Friend First | Teen First | Friend First |
| n | (207) | (197) | (305) | (99) | (156) | (51) | (149) | (48) |
| Q49A | 1.66 | $1.92^a$ | 1.76 | 1.87 | 1.68 | 1.59 | 1.84 | $2.17^c$ |
| Q49B | 1.41 | $1.62^a$ | 1.48 | $1.63^c$ | 1.39 | 1.45 | 1.56 | 1.81 |
| Q49C | 1.50 | $1.71^a$ | 1.59 | 1.63 | 1.54 | 1.39 | 1.65 | $1.88^c$ |
| Q49D | 1.36 | $1.56^a$ | 1.43 | 1.57 | 1.37 | 1.35 | 1.49 | $1.79^c$ |
| Q49E | 1.37 | $1.51^b$ | 1.37 | $1.65^a$ | 1.34 | 1.45 | 1.40 | $1.85^b$ |

Note: Superscripts denote significance level of comparisons indicated in the boxed columns.

a) $p<.01$; b) $p<.05$; c) $p<.10$

Recommendation: Eliminate the items about teen drinking and drug use but keep the items about friends' problems. If the items about the teen are kept, they should be arranged so that they are on a page following, but not facing, the friends' items.

## 7.4 Logical Consistency of Responses to Filter and Dependent Questions

The Parent Questionnaire contained a number of filter questions which were intended to either route respondents to the subsequent questions or to skip them over particular questionnaire items, depending on their response to the item. This section contains an analysis of the logical consistency of parents' responses to filter and dependent questions. Since the goal of this section is to identify filter questions that were be problematic for respondents, this analysis was limited to the 311 cases in which the respondent completed the questionnaire without the assistance of an interviewer. Only filter questions that were problematic are discussed here. If a filter question is not discussed in this section, it can be assumed that it was found to work satisfactorily. Table 7.4.1 below lists all filter and dependent questions from the 2FU Field Test Parent Questionnaire.

Table 7.4.1 - Parent Questionnaire Filter and Dependent Questions

| Filter Question | Dependent Question(s) |
|---|---|
| Q1A | Q1B |
| Q7 | Q8 |
| Q9 | Q10 |
| Q12 | Q13A, Q13B |
| Q13A | Q13B |
| Q14A | Q14B to Q16 |
| Q18 | Q18A |
| Q18A | Q19 |
| Q20 | Q21 |
| Q23 | Q24 to Q26 |
| Q39 | Q40 |
| Q44 | Q45 |
| Q55 | Q56 to Q60 |
| Q61 | Q62, Q63 |
| Q70 | Q71 to Q82 |
| Q74 | Q75 |
| Q78 | Q79 |
| Q80 | Q81, Q82 |

Analysis: **Filter question: Q14A**
**Dependent questions: Q14B to Q16**

In Q14A, respondents were asked if any language other than English was spoken in their home. Those that answered "no" were instructed to skip to Q17, while those that answered "yes" were to continue with Q14B. A total of 240 respondents indicated in Q14A that English was the only language spoken in their home. While the percentages of respondents failing to skip Q14B, Q14C, and Q14D were relatively low (3 percent failed), the percentages of respondents failing to skip Q15 and

Q16 were slightly higher. Specifically, four percent of the respondents who were supposed to skip questions 15 and 16 failed to do so.

**Recommendation:** While the skip failure rates for this question are not alarmingly high, improvements can be made. Respondents who answer no to Q14A are routed over two pages of questions. The skip failure for question 14A could be further reduced by simply including a page number reference within the skip instruction (e.g. SKIP TO Q.17 ON PAGE 10).

Analysis: **Filter question: Q55**
**Dependent questions: Q56 to Q60**

In Q55, respondents were asked if their teenager planned to continue his or her education next year. Those answering "no" were instructed to skip to Q61 and those who said their teen was planning to continue their education next year were to continue with Q56. Nineteen parents indicated that their teenager was not planning to continue his or her education during the next year. Of those nineteen, five respondents (approximately 26 percent) failed to skip each of the dependent Q56 to Q60.

**Recommendation:** It is not clear why these respondents failed to follow the skip instructions. Similar to question 14A, the skip response for Q55 skips respondents over approximately two pages of questions. Respondents might follow the skip instructions for this question more often if they are explicitly given the page number to which they are to skip.

NCES may want to reconsider the benefit of having those who answer "no" to Q55 skip Q56 to Q60. It may be of interest to researchers to be able to compare parental input for those who do not plan to continue their education directly after school to those who do plan to continue their education. By skipping these parents past questions 56 to 60, key information about their impact on their childrens' education decision making is not collected.

Analysis: **Filter question: Q70**
**Dependent questions: Q71 to Q82**

Question 70 asks respondents to identify actions they have taken to financially prepare for their teen's education. Directly beneath the stem of Q70 is a skip instruction in which respondents are to circle the number 1 and skip to question 83 if their teenager does not intend to continue his or her education. As discussed earlier in section 7.2, this imbedded skip instruction appears to have been a major contributor to nonresponse to question 70. In addition to contributing to item nonresponse, it appears that this imbedded skip instruction functioned poorly as a filter question. There were 35 respondents indicating in Q70 that their teen did not plan to continue his or her education who should have skipped to Q83. More than half of these respondents failed to skip questions 70 through 73. Skip failure rates for questions 74 through 82 were similarly high.

**Recommendation:** As discussed in section 7.2, we recommend that the skip instruction for Q70 be eliminated and a separate filter question created. Furthermore, it is strongly recommended that the skip instructions for the new filter question include a page reference.

### 7.5 Parent Data Quality Assessment

This section describes several analyses related to the quality of the parent data. These analyses involve checks on the consistency of responses to related items or of multiple response to a single item.

**Reason for Changing Schools:** Question 21 of the Parent Questionnaire asks why the student changed schools: "What was the reason for the most recent change of schools?" It lists six specific reasons, plus an "Other" option. Although the various reasons appear to be mutually exclusive, the question gives a "CIRCLE ALL THAT APPLY" instruction. Seventy-six parents chose at least one of the options listed, eight of them selecting multiple options. Five of the eight cases with multiple answers selected "Other" as one of the responses; these may involve complicated situations not adequately covered by the specific reasons listed. However, some of the other combinations appear to be contradictory. For example, one parent reported that "My teenager asked to be transferred to another school" (Option C) and "Family moved to a different location for other reasons" (Option E). It is not clear what such combinations mean.

Since analyses based on this item are likely to classify respondents into mutually exclusive categories, we recommend that the respondent choose the single best answer rather than allowing multiple responses to this item.

**Helping the Student Decide about Higher Education:** Question 57 asks about various forms of help that the parent may give to assist the student in deciding where to apply for postsecondary education ("In which of the following ways have you and/or your spouse/partner helped your teenager make decisions about where to apply for further education after high school?"). Eight possibilities are listed and the respondent is instructed to circle all that apply.

The first two options are apparently for parents who have not provided any assistance -- "Have not offered to help our teenager" (Option A) and "Offered assistance, but teenager wants to do this himself/herself" (Option B). One hundred and thirty-one respondents endorsed one of these two options. The bulk of these (93) also selected at least one of the six specific forms of help listed in the other options; in fact, 48 of these cases selected three or more forms of help. Perhaps the first two options should be dropped in favor of a final "None of the above" option; or perhaps these two options should be retained but moved to the end of the list so that parents are reminded of the different forms of help before they decide they have not given any of them.

**Saving for Post-Secondary Education:** Question 70 ("Which of the following have you or your spouse/partner done to financially prepare for your teenager's education after high school?"), Q71 ("What grade was your teenager in when you began saving for his/her education after high school?"), and Q72 ("About how much money have set aside for your teenager's future educational needs?") all concern savings by the parent for the student's postsecondary education. The three items do not always elicit consistent information.

In responding to Q70, 303 respondents indicated one or more forms of savings for education after high school; of these, 33 reported in Q71 that they had *not* yet started saving for higher education and 16 reported in Q72 that the amount they had set aside was "None". Similarly, 137 respondents did not check any of the forms of savings listed by Q70, but 17 of these indicated in item Q71 that they had begun saving for the student's postsecondary education and 19 indicated a positive amount of money set aside for this purpose in response to Q72.

Nor were answers to Q71 and Q72 always in agreement. In their answers to Q71, 119 respondents indicated that they had not yet begun saving; 29 of these nonetheless indicated some positive amount of savings for postsecondary education in Q72. Of the 235 respondents who said that they had started saving in Q71, 11 reported "None" as the amount in Q72. If we recode these items so that respondents are classified dichotomously as savers or nonsavers based on their answers to each item, the correlation between the two items is only .74.

It is not clear why these discrepancies arise; some follow-up debriefings with respondents who gave apparently inconsistent answers or cognitive interview testing may be needed during the summer instrument revision period to identify the problems with these three items.

## 7.6 Results of Parent Survey Mail Experiment

Mail surveys often suffer from low compliance rates. Because of this, techniques are frequently used to bolster the compliance rates of mail surveys. One such technique is the use of survey materials (cover letters, envelopes, etc.) that are visually appealing and formatted so they are easy to read.

The use of alternative survey materials for the Parent Survey was tested in an experiment. In this experiment, respondents were randomly assigned one of two versions of the cover letter and the envelope containing the questionnaire.

As described in Section 6.1.2, two versions of the cover letter were used. Although the content of the versions was identical, one version used a prose format similar to that used in the the Base Year parent survey, while the other used a declarative sentence style with bulleted items. Also, two types of envelopes were used for mailing. Questionnaires were sent either in standard brown NORC envelopes or in white envelopes displaying the NELS:88 logo. Parents were randomly assigned to one of the four letter version and envelope type conditions.

Logit analyses were used to determine whether the alternative survey materials reduced unit nonresponse to the parent survey. Two outcomes were studied: return of a mail questionnaire before the February 27 mail cutoff date, and return of a mail questionnaire by any date. Table 7.6.1 below displays the proportion of cases returned by mail and the proportion of cases returned by mail before the cutoff date by letter type and envelope type. As the table illustrates, the alternative survey materials did not have any effect on unit nonresponse rates.

### Table 7.6.1

### Proportion of Parents Returning Questionnaires by Mail by Letter Type and Envelope Type

| | Proportion Returned by Mail | Proportion Returned by Mail Before Cutoff Date |
|---|---|---|
| Brown envelope/ standard letter | .65 | .23 |
| Brown envelope/ bulleted letter | .66 | .20 |
| White envelope/ standard letter | .67 | .21 |
| White envelope/ bulleted letter | .64 | .24 |

It has been shown that given a questionnaire of reasonable length, more attractive survey materials can elicit greater levels of participation from respondents (Dillman, 1978). We speculate that the size of the parent questionnaire may have initially discouraged respondents from completing the parent survey. If this was the case, the impact of the alternative survey materials may have been greatly minimized.

## 7.7 Response to OMB Queries - Parent Questionnaire

In this section we report on analyses requested by the Office of Management and Budget (OMB) in correspondence relating to the clearance document.

The OMB reviewer questioned the absence of a "Not Applicable" option for certain items in Q34, Q41, and Q47, that pertain to school activities. Under the former response format, parents of dropouts or early graduates would not have been able to respond meaningfully to these items.

We suggested that we empirically assess the adequacy of two response formats, one in which a "Not Applicable (teenager currently not in school)" response column is introduced selectively, for only school-related items, and one in which it is introduced for all items. Forms A and C use the generalized response format for Q34, Q41, and Q47, and Forms B and D use the selective format.

Crosstabulation of items by response format indicates that the latter has no effect on response distribution. The "Not Applicable" option was little used, even for school-related items. Only two to three respondents used the option when it was applicable, regardless of response format; only one or two respondents circled the "Not Applicable" option for non-school items under the generalized format.

It is surprising that, of a total of 17 respondents who reported in Q18 that their teenager was no longer in school, no more than five (across both formats) selected the "Not Applicable" option for any item. Respondents may have referred to the last school their child attended when answering the school-related items.

**Recommendation:** Although the "Not Applicable" category appears to be little used by respondents who should use it based on their response to Q18, this category should be retained for the main study, given the small number of parents of dropouts or early graduates surveyed in the field test.

Though the inclusion of the "Not Applicable" response option for non-school items appears to have no effect on response distribution (respondents are correct in not using the option), selective introduction of "Not Applicable" for only school items is prudent and advisable. Selective introduction does not appear to confuse respondents and, from a formatting perspective, is not problematic.

### 7.8 Recommendations for Parent Survey

This section describes recommendations for changes in the Parent Questionnaire and data collection procedures. Sections of Chapter 6 and 7 have provided very specific suggestions for changes to individual items and procedures; the following section is devoted to general recommendations.

#### 7.8.1 Recommended Changes to the Parent Questionnaire

As noted in Section 7.2, item nonresponse is not as serious a problem for the Parent Questionnaire as it is for the Student Questionnaire. This is probably the result of having fewer items, coupled with the fact that parents were not operating under the same time constraints as students. Nonetheless, length is still a problem. Parents who think the questionnaire is too long are likely to not complete and return it at all, rather than to partially fill it out. Therefore, a shorter questionnaire may result in reduced need for follow-up and, therefore lower costs.

Generally speaking, the admonitions put forth in the recommendations for the Student Questionnaire hold for the Parent Questionnaire as well. Overcrowded format, "mark all that apply" questions, illogical format, and double-barreled response categories should be avoided. In addition, two experiments indicated that item order may affect responses, especially when the interview is conducted over the telephone (as will be the case for a significant portion of the Parent Questionnaire).

One questionnaire experiment suggest that parents' evaluations of their teenagers' high school education are affected by prior items asking about specific aspects of the high school. This led to the recommendation that the question requiring the general evaluations should precede the question requiring specific evaluations. A second questionnaire experiment suggested that responses to subsequent questions can be altered by the prior appearance of a highly sensitive question. Parents' defensiveness about delinquent behavior of their teenagers' friends increased after parents were confronted with questions about their teenager's drug and alcohol abuse.

Results from these experiments and from those included in the Student Questionnaire attest to the utility of judiciously selected and carefully constructed questionnaire experiments as an integral part of instrument development.

Other design issues and recommendations for the Parent Questionnaire follow. These suggestions do not necessarily flow out of 2FU field test experience, but are relevant to the design and planning of the main study parent component.

**Supplement the Parent Questionnaire:** A supplement to the Parent Questionnaire is recommended for the purpose of collecting data on key change measures for that segment of 2FU parents who either did not participate in, or were not eligible for, the Base Year Parent Survey. Such a supplement would extend the analytical power of the parent and student data by providing analysts with baseline measures for the parents of freshened and/or nonparticipating students. Without these supplemental data, analysts would be left with a major gap in the longitudinal dataset for this segment of parents.

The number of questionnaire items for which retrospective data could be reliably collected is limited, but is greater than could be accommodated in two-page extension of the core questionnaire. (Cost constraints dictate that the supplement to the core questionnaire be very brief.) We have identified 23 Base Year questionnaire items to consider for such a supplement. These items fall into two broad categories -- family background and financial information -- and include:

Family Background in 1988

| | |
|---|---|
| Q1b. | the amount of time the student lived with the parent |
| Q2. | number of dependents |
| Q3b. | how many siblings were in the home |
| Q4. | how many children were older than eighth grader |
| Q5. | how many children were in high school |
| Q5b. | how many had graduated from high school |
| Q6. | how many children dropped out of school before graduating |
| Q7. | marital status |
| Q.30 | highest level of education respondent completed |
| Q.31 | highest level of education spouse/partner completed |
| Q.33A | best describes current (1988) situation |
| Q.33b | held a regular job (including self-employment) |
| Q.34a | self-employed or work for someone else |
| Q.34b | categories to describe job |
| Q.36a | spouse/partner's situation |
| Q.36b | spouse/partner held regular job (including self-employment) |
| Q.37a | spouse/partner self-employed or work for someone else |
| Q.37b | categories to describe job |

Financial Information and Educational Costs

| | |
|---|---|
| Q.80 | family income from all sources |
| Q.81 | number of wage-earners in household |
| Q.82 | educational expenses for any children |
| Q.82A | total amount spent during 1987-88 school year for educational expenses |
| Q.82B | sources of money used to cover educational expenses |

It would, of course, be necessary to modify the wording of these questions so they are anchored in the appropriate time period. Because we would be collecting retrospective data for 1988, rigorous pretesting of the items through cognitive interviews to ensure their reliability is suggested.

Rather than create a separate questionnaire supplement, we recommend integrating the supplementary questions into the Parent Questionnaire as a separate section at the end of the instrument. This design is preferred over the alternative of creating two separate instruments, a core and a supplement. Two self-administered instruments are likely to appear more burdensome to a respondent than a single, slightly longer instrument. Two instruments also have the potential for increasing the nonresponse rate for one (or both) of the instruments, probably the longer core instrument.

If the supplementary questions are approved, the items will be prepared for inclusion in a two-page supplement that will appear after the last question in the core questionnaire. An instruction will be prepared to direct Base Year non-participants to complete the supplemental items.

In a similar vein, it is also worth recalling that a small number of parents completed the Base Year Parent Questionnaire, but because the sample weighting scheme required the presence of a completed Student Questionnaire, their questionnaire data was not included in the final dataset. These data are still available and the questionnaires (if the Base Year data is rereleased with the 2FU data) or selected data elements (if the supplemental Parent questions are approved) could be integrated into the final 2FU dataset and flagged accordingly. By rescuing these questionnaires we would also decrease the number of respondents requested to complete the two-page supplement in the 2FU Parent Questionnaire.

**Shorten the core Parent Questionnaire:** Parents of high school seniors are busy, just as their teenagers are; the NELS:88 2FU will be requesting information from these individuals during a period concurrent with graduation and of tremendous activity (and, in many cases, stress). As noted above in Section 7.6, we believe that Parent Survey response rates could be improved by shortening the field test questionnaire to approximate the length of the Base Year main study questionnaire. While the change in format to allow for optical scan data capture in the main study will create the perception of a shorter instrument relative to the field test version, we recommend the questionnaire be decreased in length by approximately ten questions.

### 7.8.2 Recommendations for Changes in Parent Data Collection Procedures

**Timing of the mailout:** In our proposal, we originally suggested that the mailout of the parent questionnaires should occur on a flow basis, to coincide with the administration of the student questionnaires. While this approach is both feasible and methodologically reasonable (collecting data from parents and students during the same temporally proximate period), we have revised our plan to include a single, large-scale mailout near the end of the school year, in late May or early June. Our rationale for this change is based on the following considerations.

The Parent Questionnaire includes a section of questions about the parent's plans for the teenager's postsecondary education and finances, the (parent's) decisions for which will only be known near the end of the school term or in the summer months. It is unlikely that the majority of parents will have finalized many of their decisions or crystallized their plans at an earlier point in the year. Thus, we expect to collect these data at a higher level of quality and completeness in the late spring or early summer. A single large-scale parent mailout scheduled in late May/early June will ensure that the temporal reference (the spring semester) for the twelfth grader is maintained. We anticipate that, by delaying the timing of the mailout, the follow-up effort is likely to extend through the summer months, thereby delaying the release of the final data tape. However, we believe that higher quality data for this section of the questionnaire is well worth a minor delay in the release of data.

This rationale could, of course, be expanded to a later mailing date, perhaps in September, following the HS&B example. This much later date carries with it a number of distinct disadvantages that argue against such a late mailing. First, from a methodological standpoint, other sections of the parent questionnaire may suffer from memory effects, as parents will be more dependent upon retrospective judgments for a larger number of questions. Moreover, it is likely that at least two dozen questions would have to be reworded to ensure that the questions are not misinterpreted by parents responding in the fall about their twelfth grader's experiences during the past spring semester. Second, to maintain temporal comparability in the analysis between the student and parent questionnaires, the parallel administration of both instruments is required. Certain items will require a common reference point. Third, a September mailing means that follow-up data collection activities would occur during the late fall, increasing the likelihood of encountering a substantial locating problem. Fourth, and finally, a delayed parent mailout will translate into a significant delay in the preparation of a public release data file.

**Questionnaire return:** Unlike the Base Year study, which offered parents the option to return the questionnaire via the student (or student and school) or mail it directly to NORC, 2FU parents will be asked to mail their questionnaires directly to NORC. (Two-thirds of the Base Year parent participants exercised the option of mailing their questionnaires directly to NORC.) By employing direct mail delivery, we thereby eliminate a potential source of burden on both school personnel and students.

**Nonresponse Follow-Up:** In the 2FU Parent Survey Field Test, the schedule for data collection was artificially compressed due to schedule and budget constraints. Unlike the Field Test Parent Survey procedure, the first step in main study nonresponse follow-up will involve sending a postcard to all parents thanking them for participating in the survey and urging them to return their completed questionnaires if they have not done so already. If for any reason they did not receive the questionnaire, the postcard will provide a telephone number to call toll-free in order to obtain another one. The postcard will be mailed two to three weeks after the questionnaires are sent out.

The second step in the main study nonresponse follow-up will involve a prompting call to parents who have failed to return questionnaires. We recommend, based on experience in the Base Year, that the number of telephone prompts be kept to no more than two; our past experiences suggest that a point of diminishing return is reached quickly with prompting calls. Prompting calls will begin about two to three weeks after the postcard is mailed, if a completed questionnaire has not been received. The call will be designed as a low-pressure reminder to the parent to send in a completed questionnaire. It will also serve as an opportunity to determine which parents will require remailing of the questionnaire. The telephone prompting effort will be staffed by bilingual interviewers able to prompt and conduct interviews with Spanish-speaking respondents.

## Chapter 8: School Administrator Survey Procedures

This chapter describes the data collection and preparation procedures for the school administrator component of the First Follow-Up Field Test. School Administrator Questionnaires either were mailed to the School Coordinator two weeks prior to the student survey for distribution to the principal or were mailed directly to the principal.

### 8.1 School Administrator Questionnaire: Objectives and Content

**Objectives of the School Administrator Questionnaire:** The School Administrator Questionnaire will gather descriptive data on school policies and practices that can be used to analyze the learning environment and experiences of high school students, and can help to identify features of schools that relate to positive student outcomes.

Several considerations influenced the school questionnaire design for the NELS:88 Second Follow-Up Field Test. One such consideration is that although the Base Year provided a nationally representative sample of eighth grade schools, the First and Second Follow-Ups, while preserving a representative student sample, do not constitute national probability samples of secondary schools. Hence, the sole function of the school questionnaire is to obtain contextual data for better understanding the school environment of students.

A second consideration is that the First Follow-Up school administrator questionnaire was by historical standards exceedingly long, and was perceived by school principals as excessively burdensome. (The 1FU School Questionnaire contains double the number of data elements that appeared on the NELS:88 Base Year school questionnaire, which was more nearly comparable in burden level to the HS&B Base Year instrument.) A priority of the Second Follow-Up, therefore, has been to substantially reduce the size of the School Administrator Questionnaire.

A third consideration is that the content domains for the school administrator questionnaire embrace both (largely) static variables that are unlikely to change (for example, just as a student's race is not likely to alter -- the individual and social subjectivity of racial identification does license this possibility for some individuals, but in general, race, like gender, is to be regarded as an unchanging characteristic), and dynamic variables that are more mutable -- so too do schools have certain reasonably stable characteristics over the short term (here, equaling 24 months) such that Catholic schools are unlikely to become public schools and urban schools unlikely to become rural schools between the 1990 and 1992 waves of the study. Insofaras this is true, some reduction of overall burden makes conceptual as well as logistic sense in the Second Follow-Up. In High School and Beyond's twelfth grade follow-up, the same schools were returned to, and an abbreviated school questionnaire largely updated the information obtained in the HS&B Base Year. One slight complication here is that NELS:88, while it will largely return to the same high schools in 1992, exists under the dispensation of a different sampling plan, such that in the small clusters (less than four) subsampling of dispersed students will introduce a number of new schools. One instrumentation issue that must be confronted, then, is how one may reduce overall burden, while capturing the richness of the 1FU school administrator data in instances in which a school is new to the study. Again -- as in the case of the New Student Supplement -- it would seem as though the most critical variables at risk for being lost should be specially collected from schools that are new to the study in the Second Follow-Up -- that is, First Follow-Up schools that did not provide a School Questionnaire (two percent of the sample) and schools selected for the 2FU but not the 1FU.

**Content of the School Administrator Questionnaire:** The school questionnaire is to be completed by the principal, headmaster, or other knowledgeable school administrator. The school questionnaire is intended to elicit information about the academic climate of the school by inquiring into school characteristics, policies and practices. Specific content areas include:

I. School characteristics
II. Student characteristics
III. Teaching staff characteristics
IV. School admission policies and practices
V. Grading and/or testing structure
VI. School programs
VII. School climate

### 8.2 Respondent Selection and Data Collection

The school administrator questionnaire respondent pool was defined by the distribution of NELS:88 Second Follow-Up field test cohort members; schools attended by NELS:88 students were identified by field interviewers during student tracing. (See Section 1.3 and 1.4.) The principal of each school selected for the field test was asked to complete a School Administrator Questionnaire. Principals were contacted by field interviewers during the fall activities (securing school cooperation, student tracing, and the visit to freshen and augment the sample). Principals' names were entered (and updated, as necessary) into the Survey Management System (SMS) in relation to these activities. These data were then used to identify the potential respondents for the School Administrator Survey.

### 8.2.1 Mailout and Data Collection Procedures

For 63 schools, School Administrator Questionnaire packets were mailed to the school coordinator for distribution to the principal two weeks before Survey Day. For the remaining 31 schools, the School Administrator Questionnaire was mailed directly to the principal; these 31 schools were those with the earliest scheduled Survey Day sessions. The packet contained a cover letter addressed to the principal, a School Administrator Questionnaire with ID label affixed, a NELS:88 Second Follow-Up brochure, and a postage-paid return envelope. The letter requested that the principal complete the questionnaire as soon as possible and return it directly to NORC.

Team leaders were instructed to collect the completed School Administrator Questionnaire on Survey Day if the principal had not already returned it. If the principal had not completed his or her questionnaire by Survey Day, the interviewer asked the school coordinator to remind the principal to do so as soon as possible. Often interviewers prompted the principal to complete the questionnaire in their brief meeting at the end of Survey Day to thank school officials for their participation in NELS:88 2FU.

By March 15, completed questionnaires had been received from only 30 principals. Field interviewers were provided a list of the 30 schools from which a completed questionnaire had been returned and instructed to call principals not on the list to prompt them to complete and return the instrument. These promptings were generally low pressure, especially for schools with whom we will be working in the 2FU main study.

### 8.3 Data Preparation

Data preparation encompasses all activities associated with converting information from the School Administrator Questionnaire to data files. These activities included receipt, manual editing, coding and data entry.

**Receipt Control:** School Administrator Questionnaires returned by team leaders in Survey Day materials packages were receipted upon arrival at the Lake Park Receipt Control Shop. Questionnaires returned directly to NORC by respondents were receipted on a flow basis. Receipt control clerks updated the school questionnaire disposition in the SMS from the completed instrument.

**Editing/Coding:** School Administrator Questionnaires were edited for missing items in the Edit Shop of NORC's Lake Park Facility. The focus of the edit was the assignment of standard reserve codes. Marginal indications of "Refused" or "Don't know" were coded as "7" and "8", respectively. Unmarked missing items, whether skipped legitimately or illegitimately, were marked "9" for missing, and a "6" was used to indicate an illegitimate multiple response. Open-ended questions were coded "1", to indicate that data were present. These reserve codes were preceded by a "9" for some items, depending on the number of columns corresponding to the question. For example, a missing item with response codes ranging from 00 to 12, with a marginal indication of refusal, was coded a "97".

**Data Entry:** Questionnaires were data entered at our Lake Park facility using a computer-assisted data entry (CADE) software package. Critical items were 100 percent verified (that is, re-entered independently by a different operator), with the software comparing all fields for exact matches.

No machine cleaning of the data was performed, to facilitate the identification and analysis of problems within the questionnaire.

## Chapter 9: Analysis of School Survey Data

This chapter discusses the results of an in-depth analysis of the School Administrator Questionnaire data. Response rates, overall and by school sector, are briefly examined. In subsequent sections, specific recommendations are offered for improving the effectiveness of the instrument in collecting high quality data, based on (a) item nonresponse analysis, (b) logical consistency analysis of filter and dependent questions, and (c) analyses requested by the Office of Management and Budget in correspondence relating to the clearance document. Finally, general recommendations for changes in the questionnaire are presented.

### 9.1 School Administrator Questionnaire Response Rate

School Administrator questionnaires were mailed to each of the principals of the 94 participating field test schools. A total of 65 school administrators completed and returned questionnaires for an overall response rate of 68 percent. Table 9.1.1 displays response rates by school sector type. As was the case with student participation, Catholic schools had the highest response rates. Specifically, the response rate for Catholic schools was 20 percentage points higher than that of public schools and 23 percentage points higher than the response rate for other private schools.

**Table 9.1.1 - School Response Rates by Sector**

| School Sector | Response Rate |
|---|---|
| Public | 65.7% |
| Catholic | 85.7% |
| Other Private | 62.5% |

### 9.2 Item Nonresponse Analysis

This section presents an analysis of item nonresponse for the School Administrator Questionnaire. Compared to the other NELS:88 instruments, the School Administrator Questionnaire is relatively short (a total of 24 pages, 58 questions). Because of this, it was expected that most nonresponse problems would be related to poor question formatting and unclear or ambiguous question wording, rather than respondent fatigue. As the following analysis illustrates, question format and question wording appear to be the main causes of item nonresponse for this questionnaire.

Item nonresponse rates for the school administrator questionnaire were calculated in the same way Student and Parent nonresponse rates were calculated. First, all items were dichotomized by coding valid responses as 0 and missing responses as 1. Nonresponse rates were derived by computing the mean of these dichotomous variables. The rates were adjusted to account for respondents who legitimately skipped dependent questions by removing legitimate skippers from the base of respondents used to calculate nonresponse rates for dependent questions.

In selecting items for analysis, items were first sorted by nonresponse rate in descending order. All items having an item nonresponse rate greater than 15 percent were selected for analysis. A lower cutoff was used for the School Questionnaire (a cutoff rate of 20 perent was used for the Student Questionnaire) because the administration of this questionnaire was not timed. Table 9.2.1 displays nonresponse rates for all items meeting this selection criteria.

## TABLE 9.2.1

### Item Nonresponse Rates for Analysis Selection

| ITEM | Multiple Response | Refusal | Don't Know | Missing | Total |
|---|---|---|---|---|---|
| Q3B | 0.00 | 0.00 | 0.00 | 0.18 | 0.18 |
| Q3O | 0.00 | 0.00 | 0.00 | 0.45 | 0.45 |
| Q13B | 0.00 | 0.00 | 0.00 | 0.17 | 0.17 |
| Q13C | 0.00 | 0.00 | 0.00 | 0.17 | 0.17 |
| Q13E | 0.00 | 0.00 | 0.00 | 0.57 | 0.57 |
| Q14J | 0.00 | 0.00 | 0.00 | 0.58 | 0.58 |
| Q30C | 0.00 | 0.00 | 0.00 | 0.23 | 0.23 |
| Q31G | 0.00 | 0.00 | 0.00 | 0.30 | 0.30 |
| Q31L | 0.00 | 0.00 | 0.00 | 0.48 | 0.48 |
| Q35B | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q35C | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q35D | 0.00 | 0.00 | 0.00 | 0.40 | 0.40 |
| Q36A | 0.00 | 0.00 | 0.00 | 0.23 | 0.23 |
| Q36C | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q36D | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q36E | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q36F | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q36G | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q36H | 0.00 | 0.00 | 0.00 | 0.17 | 0.17 |
| Q36I | 0.00 | 0.00 | 0.00 | 0.18 | 0.18 |
| Q36J | 0.00 | 0.00 | 0.00 | 0.73 | 0.73 |
| Q38B | 0.00 | 0.00 | 0.00 | 0.33 | 0.33 |
| Q38C | 0.00 | 0.00 | 0.00 | 0.30 | 0.30 |
| Q39B | 0.00 | 0.00 | 0.00 | 0.18 | 0.18 |
| Q39D | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q41F | 0.00 | 0.00 | 0.00 | 0.45 | 0.45 |
| Q41G | 0.00 | 0.00 | 0.00 | 0.80 | 0.80 |
| Q41SPEC | 0.00 | 0.00 | 0.00 | 0.18 | 0.18 |
| Q46A3 | 0.00 | 0.00 | 0.00 | 0.25 | 0.25 |
| Q46A4 | 0.00 | 0.00 | 0.00 | 0.25 | 0.25 |
| Q46A5 | 0.02 | 0.00 | 0.00 | 0.20 | 0.22 |
| Q46B1 | 0.00 | 0.00 | 0.00 | 0.18 | 0.18 |
| Q46B2 | 0.00 | 0.00 | 0.00 | 0.28 | 0.28 |
| Q46B3 | 0.00 | 0.00 | 0.00 | 0.28 | 0.28 |
| Q46B4 | 0.00 | 0.00 | 0.00 | 0.28 | 0.28 |
| Q46B5 | 0.02 | 0.00 | 0.00 | 0.18 | 0.20 |
| Q46B6 | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q46C2 | 0.02 | 0.00 | 0.00 | 0.15 | 0.17 |
| Q46C3 | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |

## TABLE 9.2.1 - continued
## Item Nonresponse rates for Analysis Selection

| ITEM | Multiple Response | Refusal | Don't Know | Missing | Total |
|---|---|---|---|---|---|
| Q46C4 | 0.02 | 0.00 | 0.00 | 0.28 | 0.30 |
| Q46C5 | 0.00 | 0.00 | 0.00 | 0.25 | 0.25 |
| Q46C6 | 0.02 | 0.00 | 0.00 | 0.23 | 0.25 |
| Q46D2 | 0.00 | 0.00 | 0.00 | 0.18 | 0.18 |
| Q46D3 | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q46D4 | 0.00 | 0.00 | 0.00 | 0.27 | 0.27 |
| Q46D5 | 0.00 | 0.00 | 0.00 | 0.30 | 0.30 |
| Q46D6 | 0.02 | 0.00 | 0.00 | 0.30 | 0.32 |
| Q46E3 | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q46E4 | 0.02 | 0.00 | 0.00 | 0.23 | 0.25 |
| Q46E5 | 0.00 | 0.00 | 0.00 | 0.18 | 0.18 |
| Q46E6 | 0.00 | 0.00 | 0.00 | 0.23 | 0.23 |
| Q46F3 | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q46F4 | 0.02 | 0.00 | 0.00 | 0.26 | 0.28 |
| Q46F5 | 0.02 | 0.00 | 0.00 | 0.30 | 0.32 |
| Q46F6 | 0.00 | 0.00 | 0.00 | 0.28 | 0.28 |
| Q46G1 | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q46G4 | 0.02 | 0.00 | 0.00 | 0.23 | 0.25 |
| Q46G5 | 0.02 | 0.00 | 0.00 | 0.25 | 0.27 |
| Q46G6 | 0.02 | 0.00 | 0.00 | 0.20 | 0.22 |
| Q46H4 | 0.00 | 0.00 | 0.00 | 0.27 | 0.27 |
| Q46H5 | 0.00 | 0.00 | 0.00 | 0.27 | 0.27 |
| Q46H6 | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q46I4 | 0.00 | 0.00 | 0.00 | 0.27 | 0.27 |
| Q46I5 | 0.02 | 0.00 | 0.00 | 0.21 | 0.23 |
| Q46I6 | 0.00 | 0.00 | 0.00 | 0.22 | 0.22 |
| Q46J4 | 0.00 | 0.00 | 0.00 | 0.28 | 0.28 |
| Q46J5 | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q46J6 | 0.00 | 0.00 | 0.00 | 0.20 | 0.20 |
| Q53O | 0.00 | 0.00 | 0.00 | 0.77 | 0.77 |
| Q55E | 0.00 | 0.00 | 0.00 | 0.17 | 0.17 |

Source: NELS:88 Second Follow-Up Field Test Data.

**Analysis: Q3 Indicate which of these characteristics best describes your school.**

In Q3, respondents are presented with a list of school characteristics and are asked to select the characteristics on the list that best describe their school. While the nonresponse rate for subitem A was relatively low (.07) the rates for many of the remaining items (B through O) were at or above the .15 cutoff. Many of the schools in the field test sample are comprehensive public schools. Because of this, it is likely that non responders to Q3 simply failed to circle the "no" response for subitems B through O.

**Recommendation:** Retain item as is.

**Analysis: Q13 Are students selected for a vocational education program based on the following criteria?**

Question 13 asks respondents to identify the criteria used by their school to assign students to vocational education programs. The construction of this item may be a contributing factor to the relatively high rate of nonresponse. The wording of Q13 implies that all schools assign students to vocational education programs. Instead of circling "no" for subitems A through E, respondents at schools that do not have vocational programs may have simply left the question blank.

**Recommendations:** It is recommended that a filter question "Does your school have a Vocational Education Program?", with skip instructions to guide respondents over this question if the negative response is chosen, be inserted before Q13.

**Analysis: Q30C Do department/subject area chairs or heads receive any of the following Other incentives?**

In Q30, three benefits are listed and school administrators are asked to note which of the three benefits are offered to department/subject chairs at their school. Choice C, "Other incentives" had an item nonresponse rate of 17 percent. "Incentives" may be a loaded term that has negative connotations for educators. Salesmen are given monetary incentives to sell more of a product and professional athletes are given incentives if they perform at a certain level. The term used in the context of Q30 was perhaps confusing and possibly bothersome to school administrators.

**Recommendations:** Substitute the term with a more neutral term such as "benefit" or "compensation" for the term "incentive".

**Analysis: Q31 How many full-time faculty members are in each of the departments/subject areas in your school?**

In Q31, school administrators are asked to report the number of full-time faculty members their school employs in each of 11 types of departments/subject areas; the twelfth type is "Other". Subitems G (history) and L (Other) had nonresponse rates of 30 percent and 48 percent, respectively. In reviewing the frequency distribution of responses to subitem G (history), it is evident that a relatively large percentage of high schools (30 percent) do not have a department solely devoted to history. It is likely that most of those who did not answer subitem G were also at schools not having a separate history department and, because of this, they simply neglected to circle the L response (No Department). The high nonresponse rate for the "Other" category may also be attributable to the fact that many of the nonresponders worked at schools in which all of the departments/subject areas fit into the other 11 categories and simply neglected to circle 1 for subitem I.

**Recommendation**: Retain item as is.

**Analysis: Q35 Does your school currently use any of these forms of teacher evaluation?**

Question 35 instructs respondents to indicate which (if any) of three methods of teacher evaluation are used. The item is constructed so that respondents circle either "yes" or "no" for each of the three types of evaluation. The high rate of nonresponse for Q35 is likely due to the fact that the instructions for the question are incorrect. The instructions read "CIRCLE ONE CATEGORY", when they should have read "CIRCLE ONE CATEGORY ON EACH LINE".

**Recommendation**: Correct the instructions to read: "CIRCLE ONE CATEGORY ON EACH LINE".

**Analysis: Q36 Are any of the following kinds of rewards given to teachers in your school?"**

Question 36 asks respondents to identify different types of rewards that are given to teachers. There are two fundamental problems with the construction of this question. The first problem is that the list of items presented in Q36 includes conceptually disparate elements. On one hand, the list includes things that a school administrator might manipulate in order to reward a teacher who has performed well (e.g. a teaching award, teacher is assigned to teach better students, etc.). On the other hand, the list also contains items that could be considered matters of general school policy (e.g. teachers are allowed to choose the classes they teach, teachers are given time off to attend professional workshops). The second problem with Q36 deals with subitem A, "No rewards are given". The question is constructed in such a way that respondents are to circle "No" if they do not utilize a particular reward. This format makes the inclusion of the category "No rewards are given" illogical and unnecessary.

**Recommendations**: It is strongly recommended that the conceptual intent of this question be reassessed. If this question is intended to capture information about the different types of privileges and awards that school administrators use to reward teachers, the list of items should be limited to such awards (i.e., items B, C, E, F, and G) and should not include things that are done as a matter of general school policy (i.e., items D, H, and I). Alternatively, the stem could be altered to ask if any of the items are used as rewards. It is also recommended that the choice "No rewards are given" be deleted. Respondents who do not give out rewards can so indicate by simply circling "No" for each type of reward.

**Analysis: Q38 Is the competency test a state, district, or school requirement?**

**Q39 Are the following areas covered on the competency test?**

Question 38 asks respondents whether the competency test that is administered to their students is required by the state, the district or the school; the question is in a yes-no format item. In Q39, school administrators are asked to identify which subject areas (of math, science, English, and history/social studies) are covered on the competency test, also in a yes-no format.

Item nonresponse rates for Q38B and Q38C were 33 percent and 30 percent, respectively. Perhaps many of the respondents who failed to answer these two questions work at schools where competency tests are required by the state and not by the district or the school and simply failed to circle "No" for these two items.

Nonresponse rates for Q39B (18 percent) and Q39D (20 percent) were also quite high. Many minimum competency tests only include Math and English sections. It is probable that a number of respondents from schools that administer competency tests comprised only of these two subjects failed to circle "No" for these two items, thus causing nonresponse rates to be higher than expected.

**Recommendation**: Retain items as they are.

**Analysis: Q41 When a student fails the competency test, what options are available to him or her?**

In Q41, respondents are asked to identify whether listed options are either offered, not offered or required by the school when a student fails a minimum competency examination. A major problem with Q41 is that the wording of the question's stem and the wording of the response categories are incongruous. While the stem of the question focuses on "options that are available", the response categories focus on whether options are "Offered", "Required" or "Not Offered".

Among all of the listed options, item f ("No action taken") with a nonresponse rate of 45 percent is particularly problematic. It is likely that respondents did not mark an answer for item F because the combination of this item and the question's stem is illogical. Respondents no doubt had difficulty trying to conceptualize what was meant by "No action taken" being offered, required or not offered.

**Recommendations**: We recommend the elimination of option F "No action taken". In addition to the problems cited above, Q41 could be further improved by explicitly focusing respondents on options that are available to the student that are associated with the school.

41. When a student fails the minimum competency test, which of the following options are available to the student at the school and which are required of the student?

(CIRCLE ONE CATEGORY ON EACH LINE)

| | This Option is Not Available | This Option is Available But Not Required | This is required of Students who Fail the Test |
|---|---|---|---|
| a. Retaking the test | 1 | 2 | 3 |
| b. Taking remedial classes in deficient subject areas | 1 | 2 | 3 |
| c. Completing a general competency test preparation class | 1 | 2 | 3 |
| d. Tutoring | 1 | 2 | 3 |
| e. Summer School | 1 | 2 | 3 |
| f. Other | 1 | 2 | 3 |

**Analysis: Q46 Who at your school makes each of the decisions listed below?**

Question 46 presents respondents with an immense grid on which they are to record the amount of influence six different individuals have on ten different decisions. If completed correctly, respondents are essentially answering 60 separate questions in Q46. The average nonresponse rate for each subpart of Q46 was approximately 25 percent indicating that on average, one quarter of all respondents failed to answer each of the 60 subparts.

**Recommendations**: Since such a vast amount of information is being collected with this question, it appears that little that can be done in terms of format modifications that will make the question any easier or less burdensome to answer. However, some slight modifications to the wording of the question and clarification of the questions' instructions may elicit higher levels of response, by making the question more engaging. The proposed question modification follows.

46. We are interested in how decisions are made at your school. The grid below contains 10 decisions that are often made in the course of running a school. The grid also lists 6 individuals or groups who often make these decisions. For each decision, please circle one of the following numbers for each decision maker, indicating how much influence the decision maker typically has:

0=no influence
1=minor influence etc.

**Analysis: Q55E How would you characterize your school's relationship with each of the following individuals or groups?**

In Q55E, respondents are asked to characterize their school's relationship with the "Teachers' association or union (including but not limited to the contract)". The reference to the teachers contract in this question may have served to confuse respondents. It is not clear what is intended by this reference.

**Recommendation:** We recommend that the reference to teachers contracts be dropped. Q55E would simply read "Teachers' association or union".

### 9.3 Analysis of Responses to Filter and Dependent Questions

The School Administrator questionnaire contained three filter questions; Q26, Q28 and Q37. This section contains an analysis of respondents' answers to these questions and to the related dependent questions.

**Analysis: Q26 Is your school organized into departments (or divided into subject areas)?**

In Q26, respondents were asked if their school was organized into departments. Those who answered "no" were instructed to skip to Q31 on page 11. Only three respondents indicated that their school was not divided into departments. Two of the three respondents correctly followed the skip instructions and skipped Q27 through Q30.

**Recommendation:** While the extremely small number of cases does not provide conclusive evidence about this filter question, it is recommended that the skip instructions for Q26 be retained as is.

**Analysis: Q28 Does your school formally designate a chair/heads for each department/subject area?**

Question 28 asks respondents if their school formally designates chair/heads for each department. Those who answered "no" to Q28 were instructed to skip to Q31 on page 11. Similar to Q26, only two respondents answered no to Q28. Both respondents correctly skipped Q29 and Q30. Again, it is impossible to assess how well the skip instructions for Q28 are working based on such a small number of cases.

**Analysis: Q37 Are students required to pass a minimum competency (proficiency) test(s) in order to receive a High School Diploma?**

Question 37 asks respondents if students at their school are required to pass a minimum competency test in order to graduate. Those who answered "no" were instructed to skip to Q42 on page 14. A total of 20 respondents said their students are not required to pass such a test. All 20 of these individuals correctly skipped Q37A through Q41. It is recommended that the skip instructions for Q37 be retained.

## 9.4 Responses to OMB Queries - School Administrator Questionnaire

In this section we report on several analyses requested by the Office of Management and Budget (OMB) in correspondence relating to the clearance document. OMB raised questions about each of the items or sets of items listed below. In general, the questions concerned usefulness of certain response categories and possible redundancies among certain items.

Analysis: Q6 **Approximately what percentage of your 12th grade students is in each of the following instructional programs?**

In subitem D of this question, the school administrator is asked to report the percentage of students in each of nine vocational/technical/business programs. (A similar version of this question was included in the Base Year and First Follow-Up School Questionnaires.) The OMB reviewer suggested that the field test data be used to assess the usefulness of the nine distinct vocational categories.

All of the vocational categories were used, to varying degrees. Item Q6D1, "Industrial Arts/Technology education," was the most used; 48 percent of the respondents reported that at least one percent of their 12th grade students were enrolled in such a program. "Agricultural occupations" was the least used, with 12 percent of the respondents recording data. (See Appendix E for frequencies for all response categories.)

**Recommendation:** Retain the nine distinct vocation/technical/business program categories. Greater use of the categories is expected in the main study school survey because of the greater diversity of the school population.

Analysis: Q10 **What percentage of 12th grade students do the following?**

OMB recommended that we add, as we had proposed, two additional subitems to this question and then determine their utility based on field test data. Due to an oversight, we neglected to implement this recommendation. The two additional categories would have inquired about student participation in (a) Project Talent and (b) other special programs to enhance academic skills for the transition of minority or disadvantaged students to college.

**Recommendation:** We recommend that the categories be added to the questionnaire for the main study. Two more categories should not prove burdensome to respondents, and a similar item which inquired about Project Talent and other such programs on the High School and Beyond school administrator questionnaire provided helpful in the validation of student reports.

Analysis: Q48 **Not much is known about whether and how parents of 12th grade students are involved in their children's schools and education. What proportion of parents of 12th graders in your school are involved in the following areas?**

**Item A Volunteering time frequently to help in classrooms or other school areas.**

Q50 **In the first half of the current school year, what percentage of parents or family members of 12th graders contributed volunteer time?**

OMB questioned the usefulness of Q50, in light of Q48, item A, and requested that we compare the two items to determine which is more efficient. Although Q50 and Q48A appear to request the same information from the respondent (the percentage of parents volunteering to help in school activities) the wording of the items is quite different. Q48A asks about only parents' volunteering time (whereas Q50 includes other family members), specifies the activities for which the parents volunteer their time (in Q50 the volunteer activities are left to the interpretation of the respondent), and asks about the percentage of parents who volunteer frequently.

Comparison of the data elicited by the two items reveals that there is no statistically significant correlation between responses to the items ($r=.48-p<.01$). The degree of overlap is not great ($r^2=.23$, or 23% shared variance). This suggest that the two questions are eliciting substantially (77%) unique information. The different response formats -- Q48A is categorical while Q50 is open-ended -- may also contribute to the inconsistency of the responses.

**Recommendation:** Although Q50 and item Q48A elicit different data, despite differences in question wording and response format, both seek to measure the degree of parental or family support the school receives. Each item has its faults and virtues. Question 48A has the advantages of being presented within the context of a series of items concerning parental involvement in school-related activities and of specifying the volunteer activities. At the same time, nonresponse to Q48A and the remaining subitems of the question is low, in spite of the number of items.

On the other hand, the stem of Q48A asks for a proportion, but the response categories are percentages; furthermore, the response categories are discrete rather than continuous. Q48A also requires the respondent to judge what constitutes "frequent" volunteering. Other subitems of Q48 demand that the respondent make similar judgements about the meanings of "actively", "frequently", and "regularly". The subjectivity of these judgements may adversely affect data quality.

Question 50, while it does not require that the respondent make such judgments, does not specify the activities for which the parents or family members volunteer. Its inclusion of other family members, though, is appropriate to the assessment of family support of and involvement in school activities.

If space permits the inclusion of Q48 in the main study instrument, the scope of the question should be broadened to include family members in addition to parents. "Frequently", "regularly", and "actively" should be struck from the subitems or replaced by objective qualifiers. The response categories should be continuous, as in Q10, Q15, and Q23.

If Q50 (which was fielded in the First Follow-Up School Questionnaire) is retained, we recommend that the ambiguity regarding the nature of the volunteer activities of inquiry be clarified. The revised question might read: "In the first half of the current school year, what percentage of parents or family members of 12th graders volunteered to help in classrooms or other in-school or extracurricular activities?"

**Analysis: Q46 Who at your school makes each of the decisions below?**

**Q54 How much influence do you as principal have over ...**

The OMB reviewer questioned the need for Q54, in light of Q46, and requested that we analyze and compare similar items in each. Although some subitems appear to be redundant across the two questions, close examination of the items reveals that they are not comparable. Question 46 seeks to determine how much influence various actors, including the school principal, have in specific decisions directly related to instructional practices. Q54, which was included in the First Follow-Up School Questionnaire, inquires about the principals influence over a broader range of administrative, academic, and personnel-related decisions.

To illustrate the disparities between the subitems, Q54A asks how much influence the principal has over "hiring teachers." Q46A1 asks how much influence he or she has over "establishing criteria for hiring and firing teachers." As shown in Table 9.4.1, correlation of responses to the two items reveals that, though there is a statistically significant correlation, it is slight. Only two other sets of similar items demonstrate even a weak correlation. One set consists of Q54E, which inquires about influence over "setting curricular guidelines," and Q46I, which asks about the determination of "instructional objectives of courses." The other set of items demonstrating a weak correlation comprises Q54E and Q46C1, which inquires about the influence of the prinicpal over the establishment of "policies and priorities for grouping students into classes." Though Q54E is similar to both Q46I1 and Q46C1, it is is broader in scope and encompasses the latter items.

**Recommendation:** We recommend that both questions be retained (space permitting) because of the uniqueness of the information each collects. Q46A1 through Q46J1 inquire about specific issues related to instructional practices, while Q54A through Q54L ask about similar but broader issues, in addition to inquiring about administrative decisions. (However, Q46 should be reformatted as suggested in Section 9.2, to increase item response.)

Table 9.4.1 Comparison of Responses to Questions 48A and 50

| **Correlations:** | Q54A | Q54C |
|---|---|---|
| Q46A1 | .6369<br>( 54)<br>P= .000 | .1373<br>( 54)<br>P= .322 |

| **Correlations:** | Q54K |
|---|---|
| Q46H1 | .2957<br>( 53)<br>P= .032 |

| **Correlations:** | Q46C1 | Q46D1 | Q46E1 | Q46F1 | Q46I1 | Q46G1 |
|---|---|---|---|---|---|---|
| Q54E | .3679<br>( 51)<br>P= .008 | -.0206<br>( 53)<br>P= .883 | .1739<br>( 55)<br>P= .204 | -.0879<br>( 53)<br>P= .532 | .4910<br>( 50)<br>P= .000 | .2580<br>( 47)<br>P= .080 |

| **Correlations:** | Q46J1 | Q46I1 | Q46G1 | Q46C1 |
|---|---|---|---|---|
| Q54F | .3243<br>( 50)<br>P= .022 | .3313<br>( 50)<br>P= .019 | .2576<br>( 47)<br>P= .080 | .1016<br>( 51)<br>P= .478 |

(Coefficient / (Cases) / 2-tailed Significance)

Source: NELS:88 Second Follow-Up Field Test Data

## 9.5 Recommendations for Changes to the School Survey

**Questionnaire:**

Overall, the 2FU Field Test School Administrator Questionnaire worked very well. The results of item nonresponse and filter/dependent analyses indicated that for the most part, questions and question instructions were clearly understood by respondents. However, as pointed out in section 9.2, the School Administrator Questionnaire did contain a number of problematic items. Most of the problems found were specific to a particular question and subsequently did not fit into any sort of general pattern. The reader is referred to section 9.2 for our recommendations regarding those items.

One problem involving four different items does seem worth mentioning in more detail here. A number of questions in the School Administrator Questionnaire have the "CIRCLE ONE ON EACH LINE" format. In each of these questions, respondents are presented with a grid. The rows of the grid are comprised of a list of items, while the columns of the grid are comprised of response choices (e.g. yes, no). Respondents are instructed to circle one response choice for each item on the list.

The list of response choices generally contains a category respondents can use to indicate that a particular item on the list does not apply to them (e.g., "no" or "does not apply").

Item nonresponse analysis revealed that when answering questions of this format, respondents very often do not circle one response for every item. This can have serious implications for the quality of the data collected. When a respondent fails to circle a response for every item, it is impossible to tell whether the respondent intentionally skipped the item or whether the respondent inadvertently omitted a response.

In order to reduce nonresponse, it is recommended that the format of these questions be changed. If that is not feasible, the instructions for these questions should be made more explicit. For example, on questions where the response choices are "yes" and "no", the instructions could direct the respondent to "CIRCLE EITHER YES OR NO ON EACH LINE". It is deemed that by making the question instructions more specific, nonresponse will decrease and the data collected will be of a higher quality.

**Data Collection Procedures:** No particular problems have arisen to suggest that changes in data collection procedures are needed. The sole change anticipated is motivated not by school principal behavior or responses, but rather out of an attempt to reduce the burden on school coordinators and field staff. For these reasons of burden, we recommend distributing School Administrator Questionnaires by mail directly from NORC. The package will contain a cover letter encouraging the principal to respond promptly, a NELS:88 2FU brochure, a questionnaire, and a post paid envelope to return the completed questionnaire directly to NORC. Prompting calls will be made to nonresponders from NORC's Lake Park Facility.

## Chapter 10: Preliminary Course Offerings and Enrollment Data Collection Assessment

Since the course offerings and enrollment and academic transcript components are new to the NELS:88 projects, special attention is being paid in the field test phase to ensure the success of these in the main study. The field test phase is especially important for these components, as the procedures for data collection and processing must be tested more thoroughly than in components long a part of NELS:88.

To this point in NELS:88, only student self-reports of courses taken and grades received have been collected and analyzed. The transcript study offers a method of obtaining complete course taking and grade histories for all the years of high school. This will make possible analyses of such major issues as verification and correction of student-reported grades, student curriculum exposure, and patterns and consequences of tracking.

The offerings and enrollment component will allow researchers to analyze all the courses which NELS:88 students have had the opportunity to take. If the statistical problem of approximating the selection probabilities of the NELS:88 high schools is solved, the course offerings can be weighted to reflect the course offerings of American high schools and be compared with previous national data to support trend analyses. Even should this not prove feasible, the course offerings component will furnish an extremely valuable contextual framework for understanding the learning environment of the NELS:88 student sample.

While our original plan called for only a qualitative review of materials received from field test schools, a revised strategy has been developed to more thoroughly assess the field test course offerings and enrollment data collection and processing procedures. The revised course offerings and enrollment field test, which is currently in progress, involves: (1) abstracting and entering the course offerings data for a randomly-selected subsample of 24 of the 94 total field test schools; (2) retrieving missing course offerings information and collecting enrollment statistics from those 24 schools; (3) prompting calls to school coordinators to complete and return the course offerings and enrollment data collection forms; and (4) entering the retrieved data. These course offerings and enrollment (COE) field test activities will inform plans for the 2FU main study. Because the field test for this component is yet ongoing, NORC will submit to NCES a memorandum describing final results and recommendations after completion of data collection and processing.

### 10.1 Course Offerings Data Collection Procedures

Team leaders began collecting course offerings information for the 1990-1991 academic year from all field test schools in the fall of 1990 and, when unsuccessful, continued their attempts to obtain these data throughout in-school data collection activities during spring 1991. In response to our request for a course catalog or analogous document, schools submitted a variety of documents and formats: course catalogs, course listings, student manuals and master teaching schedules. Of the 94 field test schools, eighty-six schools (91 percent) provided some form of course offerings information.

Of the 24 schools subsampled for the COE field test, 18 had provided course catalogs and six had not by the time packets were mailed on April 22, 1991. The packet mailed to the school coordinator consisted of: a letter explaining the Course Offerings and Enrollment component of the NELS:88 Second Follow-Up, general instructions for completing the forms enclosed, the Course Offerings and Enrollment Form (COEF) and the Enrollment Definition Form (EDF). We requested that these forms be completed and returned to NORC within two weeks. The mailing was scheduled at this time because a majority of schools indicated to interviewers in the spring that course enrollment figures for 1990-1991 would be available at the end of the school year.

Course Offerings and Enrollment Form: The Course Offerings and Enrollment Form (COEF) was designed to provide a standardized format for collecting course offerings information from schools in which information was either unavailable in existing school documents or could only be partially abstracted from those documents. The COEF contains 14 data elements.

The 18 schools that provided course offerings information were mailed Course Offerings and Enrollment Forms; these forms included information abstracted from the course catalog and/or related documents. Blank forms, pre-coded with the school name and NELS:88 identification number, were also provided to record additional information.

The six schools that did not provide course catalogs were mailed blank forms printed with the school name and identification number. The school coordinator was requested to complete the forms with the guidance from the general instructions and the sample COEF. If a course catalog or similar document was available, the school coordinator was asked to return the catalog to NORC in a pre-paid envelope in lieu of filling out the COEF. The school's course offerings information will be abstracted by central office staff, as were the data from the 18 processed catalogs.

Telephone calls to prompt the school coordinators to complete the documents began April 29, 1991. (Additionally, some school coordinators were asked to provide missing enrollment statistics data collection information if it had not been obtained by the interviewer during the student field period.) As of June 17, all of the schools have been prompted. Eighteen schools have returned missing course offerings information and/or enrollment figures; one school has refused to provide the information, five schools reported not receiving the packet and requested a remail of the materials; their status is pending.

### 10.2 Enrollment Data Collection Procedures

To determine when end-of-year course enrollment data would be available, team leaders completed the Procedural Checklist Form with school coordinators to determine the availability and nature of enrollment statistics. These data -- to determine how course enrollment statistics are defined by the school, the availability and nature of the data, and the best times of the school year for data collection -- were obtained during spring student data collection activities. Of the 94 schools in the total field test sample, 72 (78 percent) provided procedural information regarding collection of enrollment statistics. Overall, the majority of schools (80 percent) reported having their course enrollment data available in a computerized format, while the remaining schools (20 percent) reported that some manual preparation is involved before submission.

In Section 2 of the Course Offerings and Enrollment Form, the school coordinator was requested to record course enrollment figures for the 1990-1991 academic year (the number of students enrolled in each course the school offered for the 1990-1991 academic year) in the last column. On the Enrollment Definition Form, the school coordinators were asked to explain how the course enrollment figures are computed at their school.

If it is more convenient for the school to provide the requested enrollment data in a different format (i.e., floppy diskette, computer tape, or paper class enrollment lists), the school coordinator was made aware of the option to submit a copy of those documents instead of recording enrollment for each course on the form. Of the schools which have returned course enrollment statistics to date, the majority have submitted manually recorded enrollment figures on the COEF, even though a vast majority (80 percent) of schools reported having course enrollment in a computerized format.

### 10.3 Preliminary Recommendations for Course Offerings and Enrollment

During the prompting telephone calls, the school coordinators voiced concern about the timing, amount and type of the information requested. Several issues of importance are addressed and recommendations offered in this section.

**Burden:** School coordinators have generally complained about the level of burden imposed in the request for course offering and enrollment data. One said, "This school has been in existence for 27 years. I would have to go to the basement and hunt through many files for some of this information." When more COE field test data are available, NORC will identify high burden items of marginal analytic utility and provide NCES the list of candidates for deletion, in an effort to reduce burden on school staff.

**More Information Sooner:** School coordinators expressed need for more notice of the packet's arrival, in order to make the necessary arrangements for completion of the forms. For example, several school coordinators set up teams of teachers and guidance counselors to collect and organize the requested information, a feat of human resource management in schools that are understaffed and overburdened. NORC will distribute more information to school coordinators during the fall and winter so they may better plan for assembling the requested data.

**Appropriate Length of Time to Process Request:** The two week period to return materials to NORC (especially so close to end of the school year) was inadequate; many of the school coordinators expressed feeling pressured. To maximize cooperation, we want to give the schools enough time to collect the requested information, but at the same time we should not send the packet so early that course enrollment figures are either unavailable or the documents are put aside and forgotten. A more reasonable completion period of one month will be incorporated into main study procedures.

**Explanation of Uses of the Data:** The school coordinators questioned the purpose of the information requested. One school coordinator said, "We're in the trenches and don't see the value of this information." The majority questioned the usefulness of some data elements (e.g., when a course was first introduced or is likely to be dropped at their school or whether a course is held on or off campus). In the main study, we will include in the School Coordinator Manual an explanation of the transcript and course offerings and enrollment components of the NELS:88 2FU -- what will be requested when, options for providing the requested data, and the analytic utility of these data. As noted above, we will also explore with NCES reducing the number of data elements currently sought for each course taught at the school.

## Chapter 11: Summary of Recommendations

The objective of this Field Test Report is to document our assessments of data collection procedures and the quality of the instruments, both questionnaires and cognitive tests, for discussion as we plan the main study. The results of the NELS:88 Second Follou-Up Field Test demonstrate that the goals set for this pretest effort were attained. The free response test items -- the driving design force of the student component in many respects -- were successfully administered and analysis indicates that these items measure the same or similar domains as multiple choice questions. Sufficient numbers of observations were obtained to assess instruments and procedures and to develop recommendations for implementation in the 2FU Main Study. Qualitative data about the dropout instrument were collected through cognitive interviews and valuable insights are now available for the refinement of that questionnaire.

This section will synthesize and highlight the recommendations made in the preceding sections of the NELS:88 Second Follow-Up Field Test Report, and address several topics that are important to the design and planning of the 2FU Main Study but were not a part of the field test. The first two sections address lessons we have learned in our field test experiences, the first offering recommendations for procedural changes and the second for revisions to the questionnaires. The third section outlines current plans for the Teacher Survey. The fourth outlines plans for the Early Graduate supplemental questions to be included in the Student Questionnaire. The fifth section addresses the follow-up of sample members who were found to be ineligible in either the Base Year or First Follow-Up. The sixth discusses issues related to the School Effects Supplement (SES).

### 11.1 Summary of Recommendations for Changes in Procedures

There are few procedural changes that we recommend from our experiences in the First Follow-Up Main Study and the 2FU Field Test. Several refinements are worthy of note, however.

In our interaction with NELS:88 schools, we strongly recommend reducing the burden imposed on school coordinators and partially compensating them for their contribution to this project. Beyond the reinstatement of the honorarium, we suggest that producing more and better informational materials to engage their interest and elicit a sincere commitment to the project will be worth the time and effort in student attendance at in-school data collection sessions and in course offerings and enrollment data collection. To achieve these goals, we have urged in various sections of this report that some responsibilities be taken on by NORC directly (for example, mailing the Teacher and School Administrator Questionnaires to the respondents), more and better communication between NORC and its field representatives, on the one hand, and the school coordinators on the other (for example, an appointment with the school coordinator in a fall visit made by NORC representatives to distribute an attractive newsletter and begin to establish rapport), and a reduction in the amount of data requested from schools (that is, the elimination of some data elements of limited analytic utility from the course offerings form).

With regard to students in the longitudinal cohort, we recognize that measures must be taken to renew the commitment of these young adults at this very important and active period in their lives. We must pay particular attention to the new sample members -- those included in the sample through freshening in the 1FU and 2FU -- given their weaker propensity to participate and the possibility that respondents since the Base Year may be affected by this lack of commitment to the study. We recommend an altered schedule of communication with students (a postcard in the fall reminding them of the upcoming winter/spring data collection, a newsletter designed to be especially attractive to this age cohort two weeks before Survey Day, and a Survey Day invitation two days before the session), all designed to remind them of the importance of their participation in NELS:88.

For those members of the cohort who are no longer attending high school, we must be more sensitive to the diversity of life situations and adjust our data collection plans accordingly. Specifically, this may mean gathering data from fewer respondents in group administrations and more through one-on-one interactions (to the extent this is budgetarily feasible) and mail and telephone interviews (the latter alternatives resulting in the loss of cognitive test data.) We also strongly recommend that the fixed incentive be increased to an attractive level above minimum wage, and that interviewers be trained to react to the specific needs of the respondent by offering reimbursement for child care, transportation, etc. Such incentives are worth the expense in the long term, as fewer interviewer hours are required, on average, to complete a questionnaire and more cognitive test data are available for analysis.

In addition to relying on the postal system rather than high school seniors to deliver the Parent Questionnaire, we suggest that the timing of the data collection be changed from one coincident to the student data collection to a time in the late spring. This delay in data collection will ensure that information obtained from parents regarding plans for postsecondary education and its financing will be based on greater certainty and therefore be of better quality for analytic purposes.

## 11.2 Recommendations for Changes in the Instruments

Regarding the questionnaires, four common themes run through our recommendations. To obtain more and higher quality data, we strongly urge that these recommendations guide the revision of the survey instruments before the main study. We are keenly aware of the importance of comparability in longitudinal and cross-cohort analyses, and do not offer these suggestions without consideration to this issue. Specifically, we recommend:

. **Reducing the length of the questionnaires.** The extent of the problem varies across the instruments, with the Student Questionnaire most seriously affected and the School Administrator Questionnaire the least. We jeopardize both response rate (within budget and schedule) and quality of data by imposing such a burden on respondents.

. **Simplifying the wording and format in all questionnaires.** In many instances, the words that seek to measure important behaviors and attitudes in NELS:88 questionnaires are the words of highly educated researchers, not those of the general public and certainly not of school leavers or the educationally disadvantaged. In addition to the elitist wordings, a number of questions are too dense, byzantine in their flow, and convoluted in their logic. Since all of these instruments are meant for self-administration, they must be revised to be more "user-friendly."

. **Providing clearer instructions to the respondents.** The rationale is similar to that above. If we expect respondents to be thoughtful and thorough in their completion of the questionnaires, we must not make the task as challenging as a tax form. While that assessment is somewhat hyperbolic, response rates and data quality will be generally improved if we rely less on the respondents' motivations to do a good job filling out NELS:88 questionnaires.

. **Testing revisions through cognitive interviews and focus groups.** The revisions produced through implementation of the recommendations above should be tested through the systematic use of these techniques to ensure that the desired outcome is being achieved. No other known techniques can guide the revision process and provide the information needed for such a reasonable investment of hours and effort.

Regarding the cognitive tests, the critical issue is whether free response items should be included in 1992.

The possibility of including free response items in NELS:88 Second Follow-Up must be assessed from two vantage points. First, do these items enhance the longitudinal measurement potential of the study, within its design and testing time constraints? If so, is such an enhancement sufficiently great to justify the unbudgeted extra cost associated with scoring such items in the main study? Given the evidence of the field test and the recommendations received so far, both budgetary and measurement considerations suggest that free response items should not be included in the main study cognitive test battery.

The empirical findings of the field test, and the additional considerations upon which this recommendation rests, are discussed in detail in section 3.10 of this report. The rationale for recommending against inclusion of free response items in the 2FU main study reflects several considerations. **First, the operational conditions of creating sufficient testing time for inclusion of free response items entail dropping one of the current tests in the four-subject cognitive battery; we believe that dropping any of the four tests would be imprudent at this stage in the study.** Dropping any of the subject tests is inadvisable because (a) all four tests contribute substantially to the paramount NELS:88 testing goal of measuring the longitudinal gain of the eighth grade cohort and freshened tenth grade cohort; (b) for 1992 freshened twelfth graders, and for most 1FU nonparticipants who are surveyed in 2FU (recall that surveying as many First Follow-Up nonparticipants as possible is a priority in the Second Follow-Up), the 1992 tests will be the sole measurement point and only objective assessment of overall achievement upon completion of secondary schooling; (c) no particular test is a credible candidate for deletion. Most students continue taking social studies courses in the last two years of high school; dropping the history/citizenship/geography test would therefore seem inadvisable. Although the reading test has sometimes been regarded as dispensable at twelfth grade level, HS&B results show that, in general, reading score gain between tenth and twelfth grade was about as great as math, and that this gain was different for various groups (race, SES, school completers vs. dropouts) such that differential progress for various subpopulations intersects with primary policy concerns that NELS:88 was designed to investigate.

**Second, several distinct difficulties were identified in using free response items in the context of NELS:88 constraints and purposes.** For example, while test reliability was increased by free response items, it was increased to a lesser extent than could be achieved through adding additional multiple choice items within a like period of testing time. Given the severe constraints on NELS:88 testing time, this consideration must weigh heavily in any evaluation of the contribution of free response items to the cognitive battery.

Nonresponse in general was significantly higher in the field test for free response items than for multiple choice, while non-response on free response items was, for policy-relevant minority groups, disproportionate in comparison with nonresponse for whites.

Given the need to maximize the reliability of cross-sectional measurement in grade 12 so that a basic ability measure is achieved that will provide a benchmark for postsecondary rounds of NELS:88, given the need to maintain comparability to NLS-72 and HS&B, and given the need to give first priority to longitudinal gain measurement, the evidence seen thus far points to leaving the basic NELS:88 test design intact for this final round of cognitive measurement.

Although the field test does not appear to provide evidence that would justify recommending inclusion of free response items on the NELS:88 2FU cognitive test battery, we recommend that NCES give serious consideration to the possibility that the free response item data base assembled in the field test be shared with researchers on a restricted use or public use basis. The NELS:88 field test is based on a purposive sample the intent of which was to ensure demographic heterogeneity for evaluating test and questionnaire items. We recognize that NCES standards militate against release of databases that are grounded in non-probability samples and are therefore not generalizable for estimation purposes. Nonetheless, we feel that with sufficient caveats as to the unsuitability of the field test sample for reaching conclusions about the condition of education on a national level, the extraordinary heuristic and exploratory value of the free response data assembled in the NELS:88 2FU field test argue that serious consideration be given to the possibility that the database be made available to the psychometric community.

### 11.3 Recommendations for the Teacher Survey

This section describes the proposed design for the Second Follow-Up Teacher Survey. Presented is a discussion of the content and objectives, data collection plan including procedures for respondent selection, distribution of questionnaires, nonresponse and data retrieval follow-up, and finally, data processing of teacher questionnaires.

#### 11.3.1 Content and Objectives

The NELS:88 teacher component was designed primarily to provide teacher information that can be used to analyze the behaviors and outcomes of the student sample, including the effects of teaching on longitudinal student outcomes. The teacher-student-class linked design of this component does not provide a stand-alone analysis sample of teachers, but instead permits the direct relating of specific (rather than aggregate) teacher characteristics and practices to the characteristics and outcome measures for sampled students. The teacher questionnaire is the critical instrument for investigating the student's specific learning environment, and provides additional information on teacher background and activities, school climate, administrative context and policies. Given the curriculum-sensitivity of the cognitive tests, the teacher survey also provides a basis for assessing such class- and student-specific concepts as "opportunity to learn" in conjunction with learning outcomes in the four subject areas of the cognitive test battery.

As in the Base Year and First Follow-Up surveys, the Second Follow-Up Teacher Questionnaire will obtain information in the following four content areas:

- Teacher's assessment of the student's school-related behavior and academic performance, educational and career goals (e.g., likelihood student will go to college, student motivation, effort, absenteeism, and class participation). Respondents will be requested to complete this section with respect to the sample members they instructed for a particular subject matter.
- Information about the class the teacher taught to the sample member (e.g., track assignments, instructional methods, homework assignments, and curricular contents).
- Information about the teacher's background and activities (e.g., academic training, years of teaching experience, employment status).
- Information about the school social climate (e.g., teacher autonomy, participation in determining school policy, and relationships with the principal).

In the Second Follow-Up Teacher Questionnaire, renewed emphasis will be placed on

"Opportunity to Learn" items, which are being carefully articulated with the cognitive tests. Efforts have been made to accommodate new mathematics standards. New items have been developed to inquire into critical processes such as instructional decision-making and communications, and to ask about the quality and consequences of discussions in the course of in-service programs. Thus, for example, a communications matrix -- eliciting information about to whom the teacher talks about what key topics, how frequently and where -- should prove to be an unusually comprehensive and powerful item for capturing a phenomenon that has been little studied at such depth in previous large-scale quantitative research.

Items that did not work well in the First Follow-Up will be deleted or suitably revised. First Follow-Up school-level items that do not involve critical change measures should generally not be repeated on the Second Follow-Up teacher questionnaire, since we are returning again to largely the same school sample in 1992 that was surveyed in 1990. Given that there has been no field test of the teacher component, it is epecially critical that the teacher questionnaire be thoroughly assessed through such qualitative techniques as cognitive interviewing, or through focus groups and iterative pretests.

The target completion time for the teacher questionnaire is thirty minutes. To reach this completion time, a considerable reduction in size of the First Follow-Up instrument will be required. By reducing overall teacher burden, we expect to achieve far better rates of unit response and moderately improved levels of item response than were achieved in the First Follow-Up. (The thirty-minute Base Year Teacher Questionnaire was completed by 92 percent of teachers; the sixty-minute 1FU instrument was completed by 84 percent of teachers.)

### 11.3.2 Data Collection Plan

The Second Follow-Up Teacher Survey will consist of a thirty minute self-administered questionnaire mailed directly to 4,000 core teachers and 2,500 School Effects Study (SES) teachers. The expected completion rate is 93 percent, yielding 3,720 core teacher questionnaires and 2,325 SES teacher questionnaires. It is anticipated that 60 percent of the questionnaires will be completed by mail with telephone follow-up, and 40 percent completed by telephone administration.

### 11.3.3 Procedure for Selection of Teachers

All full- and part-time teachers who will be teaching classes in mathematics and science (and additionally English and social studies, for students in SES schools only) to eligible NELS:88 sample members in the fall term of 1991 will be included in the NELS:88 universe of 12th grade teachers.

Selection of teacher respondents for each student will be based on the assignment of up to two curriculum areas per school -- mathematics and science. Substitutions of the curriculum areas English and social studies will occur **only** for SES teachers. Thus, for example, if a core student is not enrolled in either mathematics or science, teacher data will not be collected. It should be noted that this selection of teachers departs from previous NELS:88 teacher surveys in the following two ways: (1) substitutions in all subject areas were allowed for teachers of core students; and (2) BY subject combinations were repeated in the 1FU. The primary reason for this change in design is because funding is available only for the mathematics and science portion, which NSF has agreed to sponsor.

Once the sample of 12th grade schools and students has been specified, team leaders will complete and return to NORC a Class Schedule Form (CSF) indicating for each sampled student the name of the student's fall term teacher(s) in mathematics, science or both. For students in SES schools only, names of teacher(s) will be collected for English or social studies if the student is not enrolled in a mathematics or science class. Each teacher listed on the CSF will be asked to complete a teacher questionnaire.

The identification by teacher of the specific class that each student attended will not be collected by the team leaders as in the 1FU. Rather, the class link will be established by directly asking the teacher to specify in the questionnaire the particular class attended by each student. In the 1FU, the CSF was the mechanism by which teacher ratings of students and descriptions of curriculum and practices were linked to individual students. This information was then used to produce a list of classes for which each teacher provided descriptive information in Part II of the questionnaire. The collection of this information by team leaders via the CSF was extremely complicated and error prone. Thus, a more simplified approach for establishing this link -- directly asking the teacher -- is proposed for the 2FU Teacher Survey.

### 11.3.4 Procedures for Distribution and Collection of Teacher Questionnaires

During the spring of 1992, each selected teacher will be mailed a questionnaire packet directly from NORC. This packet will contain the Teacher Questionnaire, the list of NELS:88 students thought to be in one of the teacher's classes, and a cover letter explaining how teachers were chosen and requesting participation in the study. The questionnaire administration schedule will allow approximately three weeks for teachers to return the completed questionnaire directly to NORC. During the fourth week, a prompt/thank-you postcard will be sent to all selected teachers. This postcard will serve a dual purpose; a reminder for those who have not yet returned their questionnaire and a thank you for those who have participated. By the sixth week of the administration schedule, all nonrespondents will be mailed another questionnaire with a cover letter which makes a stronger plea for cooperation. Further follow-up of nonrespondents is discussed below.

It should be noted that the design described above proposes a deviation from the past questionnaire distribution procedure of mailing the questionnaires to the school coordinator for dissemination to selected teachers. This departure is an attempt to reduce the overall burden placed upon the school coordinators. Moreover, direct mailing to the teachers is perceived as favorable and affords the opportunity to re-establish a direct line of communication and rapport between the teacher and NORC/NELS:88.

### 11.3.5 Nonresponse and Data Retrieval Follow-Up Procedures

Beginning the ninth week of the administration schedule, telephone follow-up calls will be made to nonresponding teachers. All telephone calls will be made from NORC's Telephone Survey Center by trained telephone interviewers. Telephone interviewers will prompt for return of completed self-administered questionnaires and, when necessary, administer the questionnaire by telephone. Unlike the 1FU Teacher Survey, teachers in the 2FU survey will not receive an honorarium. Thus, a higher level of effort for gaining survey cooperation vis-a-vis more prompting calls including telephone administration is anticipated. Given the higher level of effort expected, the reduction in burden placed upon the teachers is deemed to offset potential reluctance to participate.

Following receipt of completed questionnaires into the Survey Management System (SMS), questionnaires will be edited for missing critical items. If a teacher questionnaire contains at least one missing critical item, a data retrieval telephone call will made to the teacher respondent. Data retrieval telephone calls will be made by trained telephone interviewers in the Telephone Survey Center. It is expected that 50 percent of the mail completed questionnaires and 20 percent of the telephone completed questionnaires, or a total of 1,344 questionnaires (36 percent), will require critical item retrieval. Retrieved data will be recorded on the hardcopy questionnaire prior to shipment to the optical scanning contractor so that all data are scanned and included in the data file.

### 11.3.6 Data Processing

NORC proposes optical scanning as the method for data capture of the teacher questionnaire. When the scanner subcontractor completes data conversion and a raw data tape is returned to NORC, machine-editing will be performed. This machine-editing program will perform range and code checking, and additional inter-item consistency checks for selected items. Violated skip patterns will be cleaned to consistency in accordance with predetermined standards.

### 11.3.7 Who Should be Surveyed: Autumn Term Teacher or Spring Term Teacher?

We recommend that serious consideration be given to the issue of which teacher--autumn term or spring term--be asked to complete the NELS:88 2FU Teacher Questionnaire.

To the extent that a student has, as will often be the case, the same teacher in autumn and spring, the decision as to autumn or spring teacher is without consequence. If overwhelmingly students were in year-long courses with the same teacher, either spring or the autumn option could be pursued without affecting the overall study design. One could then decide on operational and logistic grounds--if it is easier to identify the teacher in the autumn visit to the school for sample freshening, then this would be the preferred assignment point. Otherwise, assignment of the teacher should wait until spring term course information becomes available. As a matter of fact, the autumn, when survey personnel will be at the school for sampling activities, is in terms of survey costs, school burden, and operational efficiency, the better time to secure the class schedule information that will form the basis of the teacher-student linkage.

We assume, however, that significant numbers of students will have different teachers in the spring than they had in the fall. If this is so, then it is not a matter of conceptual indifference which term's teacher is selected. One must choose which teacher--fall or spring--should report on the student, and make this choice primarily on the basis of its implications for the study design.

The case for the spring term teacher can be fairly simply made, and provided a convincing rationale for the practice of both the Base Year and the First Follow-Up. Unity of temporal reference in student, teacher, and parent reports is highly desirable. Since the basic analytic unit in NELS:88 is the student, contextual data from teachers and other sources should normally be obtained for a like point in time in order to maximize comparability. If the student is surveyed spring term, then the spring term teacher should be surveyed. Additionally, even a modestly retrospective report is normally to be avoided if one can obtain a meaningful report in the present moment. Particularly this is so when one is making a real demand on memory, as one is when one asks teachers to supply ratings, since high school teachers typically deal with many students, and all the more of them if there is change in their classes between semesters. Finally, all things being equal, cross-wave consistency is a value in itself in a longitudinal study. Unless there are compelling reasons to depart from precedent, there is a presumption in favor of the status quo.

The case for the autumn term teacher may be made with reference to the changing circumstances of the survey. One important aspect of this is that while the field period has always been January or February through June, there have been significant shifts in school preferences for when survey activities should take place--or when they could take place. Building the Base Year sample was an arduous task that took a considerable time period to accomplish. Perforce, few survey sessions could take place in February, and the bulk of the sessions occurred in April, with many transacted in May and even early June. The school sample of the First Follow-Up, however, could be identified many months in advance, and cooperation for the study was secured a year ahead of data collection sessions, giving more flexibility in scheduling. Survey activities were quite heavy in late January and February, with the overwhelming majority of sessions completed by the latter part of March. Given the experience of the Second Follow-Up field test and of main study contacting to date, schools have generally expressed strong preferences for holding survey sessions as early as possible in 1992. We anticipate that most Second Follow-Up survey sessions will be scheduled for January and February of 1992. This will normally be the very beginning of the spring term, though on some school calendars, survey sessions would fall at the end of the autumn term. Such a time frame would benefit the study--if NAEP-NELS:88 equating in mathematics results is to take place, it is desirable that testing take place within the same time frame--NAEP will be in the field in January and February of 1992. It may also be beneficial from the point of view of participation and motivation on the part of students--as graduation approaches, disengagement increases. There is some basis, then, for wanting to embrace, rather than resist, this tendency toward earlier survey sessions.

To be sure, there are conceptual arguments for trying to hold survey sessions in May and June--since postsecondary plans and decisions will be better known at that date. Practically, however, we see great risks and limited prospects of success in urging schools to accept survey sessions at the very end of the term, and the lack of opportunity for make-up sessions and other follow-up provides another strong argument against a late spring survey session approach. In order to gather further information about the transition to postsecondary education or work at a more optimal time point, we have recommended that the parent survey take place at a later time--not, as in the Base Year, in parallel with the school sessions, in January and February but more toward June.

If we then confidently expect the bulk of survey activity to occur at the end of the first term or quite early in the second term of senior year, the teacher survey poses a dilemma for spring-term teachers that it does not for instructors in the autumn term. Undoubtedly one of the most important features of the NELS:88 teacher survey is that of the teacher ratings of individual NELS:88 students. While a principal argument in the Base Year against surveying the autumn term teacher was that considerable time would have elapsed since the student and teacher had last been in regular contact, the likely timing of 1992 survey sessions suggests that the memory effects/retrospection problem should not be a significant factor if the fall term teacher is surveyed. With a full term of quite recent contact, the autumn teacher should be able to provide sound student ratings in January-February 1992, while the spring term teacher will have had rather little contact (in some cases no contact--this is particularly likely in schools whose autumn term ends in late January or in February) with the student at the time of the student survey.

The problem of lack of student-teacher contact early in the spring term was recognized in the First Follow-Up although we anticipate that it will be significantly more serious (sessions yet earlier in the term) in the Second Follow-Up. The strategy for handling this problem was to extend the period of student-teacher exposure in the spring term by having the teacher survey lag behind the student. While this helps to solve the student ratings problem, it may be thought to do so to the detriment of curriculum exposure items that best may be answered if testing and the teacher survey take place simultaneously, and to pose some significant operational difficulties as well.

Student cognitive tests and teacher questionnaires ideally should be completed at the same point in time. Since the tests are curriculum-sensitive, and the teacher questionnaire will contain curriculum content and exposure/opportunity to learn items, articulation between test content and teacher-reported curriculum content must be achieved through a common temporal anchor point. To be sure, the spring term teacher will probably know what the autumn term teacher covered (or was supposed to cover), and the autumn term teacher will know what is to be covered in the beginning of the spring term. But there is a further difficulty in supplying a proper anchor relative to the test administration date. If response categories for a curriculum topic instanced in a teacher questionnaire item will include such options as "already covered this year" and "to be covered this year" the line between past and future should be drawn by the test-taking date; accuracy of response will be maximized by having a like completion point for student test and teacher questionnaire. If we test in January and February there is some prima facie desirability to surveying teachers then too, rather than waiting until say May. Also, the closer one gets to the close of school, the less time one has to follow up on nonresponding teachers, and the less time one has to retrieve missing critical data. Teachers disperse in the summer; they are hard to contact. The costs of locating them after schools close is high, and the results uncertain. Surveying teachers at the start of the spring term is more efficient than surveying them at the end of the term. If students have had minimal exposure to spring term teachers at that juncture, then perhaps the autumn term teacher should be surveyed.

As a pendant to this discussion, it may be worth noting that senior year is different from tenth grade and eighth grade in a respect that is importantly relevant here. As one approaches graduation, notoriously, there is disengagement from school tasks. Some schools do not require second semester seniors to take final exams. Even where schools do require seniors to take exams, the second semester of senior year typically represents a period of less avid pursuit of coursework. Teacher ratings of students compiled by spring term teachers, especially in the late spring, therefore represent an anomalous assessment point, a point of highly atypical behavior and academic performance. If the student is well known to the teacher, the teacher can take a long, retrospective view and take "senior slump" into account in completing the ratings. But if this wider context is lacking, the rater can only speculate on the degree to which observed behavior is typical of the student or not.

The answer to the question of which teacher, first term or second, should respond to the teacher questionnaire may be closely related to two other issues that it would be very helpful to resolve, and that have something of a life of their own. These issues are: is it desirable for the cognitive tests of students and the teacher survey to be administered at the same time point? In cases where a student may change teachers between fall and spring term, is it desirable to maximize the reference period of teacher-student contact upon which the teacher ratings of individual students will be based? What is the downward limit, the minimum duration, of that contact that we should be willing to live with? If the spring teacher is the target of the teacher questionnaire (again, this issue arises only when the spring and autumn teachers are different), can the goal of coincident test and teacher survey administration be harmonized with the potentially conflicting goal of maximizing teacher-student contact before ratings of the student are requested?

There are clearly scientific pros and cons to either approach. We recommend that the issue of autumn teacher versus spring teacher, and all the considerations that attend upon that choice, be given serious consideration.

### 11.4 Recommendations for the Early Graduate Supplement

Assuming that members of the eighth grade cohort (augmented by First Follow-Up freshened students) become early graduates at the same rate as did members of the HS&B sophomore cohort (4.8 percent), then the Second Follow-Up core sample will contain approximately 950 early graduates. Early graduates are to be retained in the Second Follow-Up sample with certainty.

Early graduates will initially be identified through school records during the proposed roster update during the autumn of 1991. An additional roster update of early graduates will be accomplished by telephone prior to Survey Day. Early graduates are asked to complete the student questionnaire, thinking back to when they were in school to provide answers about the high school last attended. Supplemental questions will be posed to the early graduates. (See Appendix E for proposed supplemental questions.) In HS&B, these questions were contained in a separate booklet. Another option, however, is to produce the early graduate supplement as a section of the main questionnaire. Reducing the number of forms that must be employed, and later, matched, has some benefit at various stages in the project, particularly in terms of simplifying mailouts, survey administration, receipt control, and document reconciliation. On the other hand, including the special questions for early graduates within one document entails more printed pages in a form that will be fully used by only five percent of the sample.

In terms of the content of the early graduate supplement, we plan to follow closely the general model provided by the HS&B early graduate supplement by inquiring into the motives for graduating early, the participants in and timing of this decision, and the extra efforts required to accomplish this end. The supplement would also determine whether the respondent has entered the labor force or is in postsecondary schooling. Although the parents of early graduates will also be surveyed, we do not propose to include supplemental questions for them. Nor would teacher data be obtained for this population; although if the autum teacher is asked to complete the teacher questionnaire rather than the spring term teacher, some number of early graduates could be linked to teacher reports.

## 11.5 Further Plans in the NELS:88 Second Follow-Up for Addressing the Problem of Excluded Students

A major component of the Second Follow-Up main study is the followback of First Follow-Up ineligible students. Though not implemented as part of the NELS:88 Second Follow-Up field test, the design and plans for this component should be reviewed as part of the field test reassessment of the overall Second Follow-Up design and as part of the process of making final recommendations for how to proceed in the main study.

**Background - Ineligible Students in the Base Year:** In the NELS:88 Base Year students were sampled through a two-stage process. First, stratified random sampling and school contacting resulted in the identification of the school sample; second, students were randomly selected (with oversampling of Hispanics and Asians) from within cooperating schools.

Prior to student selection, school coordinators -- members of the school staff, typically an assistant principal or guidance counselor who acted as liaison between the school and the study -- were asked to examine the sampling roster and assess each student's eligibility. Excluded from the target population of eighth grade students within those schools were individuals with mental handicaps that precluded meaningful assessment under standard testing conditions, students whose command of the English language was not sufficient for understanding the cognitive tests and other survey forms, and students with physical or emotional problems that would make it unduly difficult for them to participate in the survey. To better understand how excluding students with mental handicaps, language barriers, and severe physical and emotional problems affects population inferences, data were obtained during Base Year sampling on the numbers of students excluded as a result of these restrictions. (For further details on base year ineligibility rules, see the NELS:88 Base Year Sample Design Report: Spencer, Frankel, Ingels, Rasinski and Tourangeau, 1990).

The total eighth grade enrollment for the NELS:88 sample of schools was 202,996. Of these students, 10,853 were excluded owing to limitations in their language proficiency or to mental or physical disabilities. Thus 5.35 percent of the potential student sample (the students enrolled in the eighth grade in the 1,052 NELS:88 schools from which usable student data were obtained) were excluded. Less than one half of one percent of the potential sample was excluded for reasons of physical or emotional disability (.41 percent), but 3.04 percent were excluded for reasons of mental disability, and 1.90 percent because of limitations in English proficiency. Thus about 57 percent of the excluded students were ineligible by reason of mental disability, about 35 percent owing to language problems, and less than 8 percent because of physical or emotional disabilities.

The overall exclusion rate, while not surprising (in comparison, NAEP excluded about the same proportion -- 5.3 percent of the sample as compared to almost 5.4 percent for NELS -- from its 1988 eighth grade assessment), is particularly disturbing when one considers the disproportionate representation among the excluded of policy-relevant minorities (blacks and Hispanics) and, by definition, of special policy-relevant classes of students receiving or requiring special services (for example, Limited English Proficiency students, and students with mental or physical disabilities).

**The Base Year Ineligibles Study in the First Follow-Up:** For two primary reasons it was decided in the First Follow-Up to conduct a followback study of the Base Year ineligibles. First, the current characteristics and probable future educational outcomes for the excluded groups may depart from the national norm, and to the extent that this is so, estimation will be affected accordingly. (For example, if the overall propensity to drop out between the eighth and tenth grades is four times as high for excluded students as for nonexcluded students, the dropout figures derivable from the NELS:88 First Follow-Up (1990) study would underestimate early dropouts by about twenty percent.)

In a school-based longitudinal survey such as NELS:88, excluded students carry a second implication for future waves. Unless the eligibility of baseline ineligibles is reassessed in 1990 and 1992, and those whose eligibility status has changed are given a chance of re-selection into the sample, the 1990 student sample will depart from a fully representative sample of tenth graders, and the 1992 sample will not constitute a cross-sectionally valid sample of twelfth graders. Sample freshening to give a 1990 and 1992 chance of selection to those who were not in eighth grade in 1987-88 is a necessary, but not sufficient condition of the representativeness of the tenth and twelfth grade samples. Excluded students must also be given a chance of inclusion, if their eligibility status has changed.

A substantial subsample (approximately 600) of the Base Year Ineligibles was therefore followed in the NELS:88 First Follow-Up, to reassess eligibility status. Those who were found to now be eligible for the First Follow-Up (either because of change, inaccurate past assessment, or conformity to the liberalized language eligibility criteria of the First Follow-Up) were also surveyed using the Student Questionnaire (if in school) or Dropout Questionnaire (if out of school). For those classified as remaining ineligible, information was gathered about their demographic characteristics and school enrollment status. Data on persistence in school obtained from this subsample are being used to derive an adjustment factor for national estimates of the eighth grade cohort's dropout rates between spring of 1988 and spring of 1990. (For further details of the NELS:88 1FU Base Year Ineligibles Study see Ingels, 1991.)

**The First Follow-Up Ineligibles Study in the NELS:88 Second Follow-Up:** In the 2FU main study, Base Year Ineligibles who were found to be First Follow-Up eligible -- whether dropouts or students -- will be folded into the core sample.

The Base Year Ineligibles who were found to be still ineligible in the NELS:88 1FU will constitute the bulk of the students in the 1992 followback of First Follow-Up Ineligibles. Two additional groups of students, however, will also be included in this component. First, a small number of 1FU students selected for freshening were declared ineligible. These students should be included in the followback study of excluded students. Second, a quite small number of students who were eligible for the Base Year became ineligible for the First Follow-Up. In rare cases, mentally or physically incapacitating events will befall a sample member who was able to participate in a prior round. These students too should be included in the 1992 followback.

The Study of 1FU ineligible students in the NELS:88 2FU will pursue essentially the same objectives as informed the 1FU's followback of Base Year ineligible students. Since the competence of any of these classes of excluded students may change between 1990 and 1992, their eligibility status should be reassessed. In addition, complete school enrollment status information should be obtained, as well as confirmation of their basic demographic characteristics.

**Recommendation**: The general recommendation that we would like to make is that NELS:88 continue in the direction pointed by the First Follow-Up's followback of excluded Base Year students and by the universal enrollment status screening adopted in the 1FU dropout survey.

The traditional -- HS&B and NELS:88 Base Year -- definition of survey participation is completion of the Student Questionnaire (minimum case), or the Student Questionnaire and the cognitive test battery (maximum case). Nonrespondents -- those for whom there is no completed questionnaire -- receive no final (nonresponse-adjusted) weight and do not appear on the final data file.

However, an alternative approach is to acknowledge a second level of presence in the study, based on whether school enrollment status information and the most basic sociodemographic classification variables can be obtained. Particularly for generation of school retention and dropout statistics, and in order to statistically accommodate both students who are incapable of participation in the sense of test and questionnaire completion and students who are capable but did not participate, we recommend that an attempt be made through the school or other proxy sources to obtain basic demographic and school persistence data for nonparticipants and ineligibles. A special weight can then be created to reflect this expanded definition of the "participating" survey population, as is currently being done in the First Follow-Up.

### 11.6 Issues Related to the School Effects Supplement

A School Effects Supplement to the NELS:88 First Follow-Up will be continued in the NELS:88 Second Follow-Up and is intended to examine the role of school structural and management characteristics, practices, and "climate" in influencing student outcomes and "school effectiveness". The foci of this supplement are school-level characteristics, aggregate student-level characteristics, and aggregate teacher characteristics that are presumed to lead to increased school effectiveness (for example, higher student achievement levels and persistence).

This supplement will permit NELS:88 to be used for the kinds of longitudinal school effectiveness analyses that were underwritten by the HS&B sophomore cohort for 1980-82 (with the addition of school organizational data from the Administrator and Teacher Survey HS&B subsample in 1984), despite the fundamental design difference between the two studies: the HS&B student sample largely remained in the same schools between the tenth grade beginning point of the survey and twelfth grade; NELS:88 sample members, who began the survey as eighth graders, overwhelmingly changed schools.

Under the original design adopted for the study -- to investigate the individual correlates of learning by following a representative sample of students over time, regardless of the high schools to which they disperse -- without a special school effects supplement, NELS:88, unlike HS&B, would lack (a) a probability sample of schools (or its functional equivalent -- specially estimated school weights), (b) a within-school representative student sample at tenth grade, (c) a sufficiently large number of students per school to permit use of analytic techniques such as hierarchial linear modeling, and (d) a large enough sample of students and teachers to facilitate investigation of the internal culture and organization of schools. In order to address the problem of the absence of a probability sample of tenth grade schools, it was decided to explore means to approximate the selection probabilities of the NELS:88 high schools (for details of methods for calculating selection probabilities for a non-probability sample of schools, see Spencer and Foran, 1991). In order to provide a representative and sufficiently numerous within-school student sample to permit extimation of within-school relationships and estimate mean achievement, socioeconomic composition, and other school-level charateristics, the student (and through the student, teacher) sample in 270 urban and suburban public and private schools in the 30 largest Metropolitan Statistical Areas (MSAs) was augmented (to achieve an average density of about 30 students) in the First Follow-Up. We shall return to these same schools in the NELS:88 Second Follow-Up.

In some fundamental respects, the NELS:88 school effectiveness study constitutes an advance on the HS&B design, insofar as HS&B was able to measure the impact of secondary schooling for only a two-year period. Given the contribution of eighth grade data to the School Effects Supplement, the full impact of secondary schooling, and the consequences of critical tracking decisions made just prior to entry into high school may more fully be assessed. Moreover, the tenth and twelfth grade teacher data for NELS:88 will be far richer than the data collected through the "teacher comment form" in HS&B; to the degree that the teacher instrument is the key to defining the student's specific learning environment and experiences, school effectiveness issues may more deeply be explored in NELS:88 than was possible in HS&B. Finally, although a major school effectiveness comparison drawn from HS&B data was the contrast between public and private schools, the non-Catholic private school sample of HS&B was thinly represented. Even within the limitation of the school effects subsample to the thirty largest MSAs, the NELS:88 private school sample offers a more robust basis for public-private school effectiveness comparisons.

Nevertheless, with quadruple the number of schools and students that are contained in the NELS:88 supplement, the HS&B dataset possesses unmatched statistical power and generalizability. No rural schools are included in the school effects subsample of NELS:88 high schools. Suitably weighted, the estimates of school effects obtained from the NELS:88 subsample may be used to make inferences about the school effects for all high schools in the 30 largest MSAs -- but not for schools outside these MSAs.

Our recommendations concerning the School Effects Supplement embrace two kinds of concerns. First, that the design for implementation of the supplement in the Second Follow-Up be assessed by independent researchers interested in using the school effectiveness data. As a last-hour add-on to the First Follow-Up, features of the design pertaining to the Second Follow-Up did not receive exhaustive attention. For example, it is our assumption that the primary intent of the supplement is to draw on longitudinal change measurements in order to assess school effectiveness between 1990 and 1992 (and, employing a portion of the student sample, for a yet longer time period). Nevertheless, it is unclear to what extent it would be desirable to maintain a within-school representative student sample in school effects subsample schools in 1992. A representative within-school twelfth grade student sample was not realized in HS&B in 1982 -- nor was the sample freshened to provide an overall representative twelfth grade cross-section. Nevertheless, by freshening on this sample and accommodating the pattern of transfer into and out of the 270 NELS:88 school effects subsample schools, a representative within-school 1992 student sample could be obtained. This issue, and others as well, could profitably be examined before the school effects design is unalterably fixed.

The second kind of recommendation that we would make is less a matter of statistical design than of operational implementation. The School Effects Supplement in the First Follow-Up was not universally well received by the subsample of schools that were drawn into it. While the research rationale for the supplement was convincing, neither districts nor schools always responded positively to the relatively sudden and significant upward definition of their level of participation in the First Follow-Up. A school, for example, that had cheerfully complied with a request for the involvement of three students and two teachers, might feel taken aback by a request for a complete annotated roster from which another thirty students could be drawn, to be followed by a request for an exhaustive teacher-student-class schedule matrix, from which a substantial teacher sample would be drawn. Given that schools felt that they had negotiated the terms of their contract with the study earlier, the redefinition of their burden levels -- accompanied by the urgency of requests for turnaround of rosters and intricate class schedule forms for sampling teachers -- sometimes generated ill-feeling and dampened the spirit of cooperation. While the NELS:88 1FU achieved an overall school cooperation rate of nearly 99 percent, just 70 percent of schools agreed to the School Effects Supplement, even though this subsample was drawn from schools already committed to the wider study. Among schools that did agree, complaints about the burden imposed by the supplement were commonly voiced, and augmented students cooperated at a rate dramatically lower than that of their longitudinal cohort peers.

We expect to return to the School Effects Supplement districts and schools with the advantage of time -- advance warning of what is required is an essential ingredient to buying commitment. We also expect to return to these schools with evidence that despite the addition of components such as offerings and enrollments and transcripts, the Second Follow-Up has made a determined effort to reduce the overall level of school burden encountered in First Follow-Up. In addition, special efforts will be required to motivate students who are not members of the longitudinal cohort and whose identification with the study is less certain.

# REFERENCES

Blalock, H.M. (1979). Social Statistics. New York: McGraw Hill.

Burton, S., and Blair, E. (1991). Behavioral frequency questions. Public Opinion Quarterly, 51, 50-79.

California State Department of Education. (1989). A Question of Thinking: A First Look at Students' Performance on Open-ended Questions in Mathematics. Sacramento: California State Department of Education.

Dillman, D.A. (1978). Mail and telephone surveys: The total design method. New York: Wiley.

Dunteman, G., Peng, S., and Holt, M. (1974). Composite Score Analyses: Ability Index, SES Index, some Psychological and Educational Construct Scales. Research Triangle Institute, Research Triangle Park, N.C.

Fetters, W., Stowe, P., Owings, J. (1984). Quality of Responses of High School Students to HS&B Questionnaire Items. Washington, D.C.: National Center for Education Research.

Ghiselli, E.E., Campbell, J.P., and Zedeck, S. Measurement Theory for the Behavioral Sciences. San Francisco: W.H. Freeman.

Ingels, S.J. (1991). The Problem of Excluded Baseline Students in a Longitudinal Study: Correcting National Dropout Estimates and Accommodating Eligibility Change Over Time. Paper presented at the American Educational Research Association Annual Meeting, Chicago, Illinois, April 6, 1991.

Ingels, S.J. and Scott, L.A. National Dropout Statistics from a Longitudinal Cohort Perspective. Paper presented at the American Educational Research Association Annual Meeting, Chicago, Illinois, April 3, 1991.

Ingels, S.J., et al. (1987). Field Test Report: National Education Longitudinal Study of 1988. University of Chicago: National Opinion Research Center.

Jobe, J.B., and Mingay, D.J. (1989). Cognitive research improves questionnaires. American Journal of Public Health, 79, 1053-1055.

Krosnick, J.A., and Alwin, D.F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. Public Opinion Quarterly, 51, 201-219.

Rock, D.A., and Pollack, J.M. (1991). Psychometric Report for the NELS:88 Base Year Test Battery. Washington, D.C.: National Center for Education Statistics. (NCES 91-468).

Romberg, T.A. (1982). The Development and Validation of a Set of Mathematical Problem-Solving Superitems. University of Wisconsin: Wisconsin Center for Education Research.

Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80, 1-28.

Royston, P., Bercini, M., Sirken, M., and Mingay, D. (1986). Questionnaire design research laboratory. Proceedings of the Section on Survey Methods Research, American Statistical

Association, 703-707.

Schuman, H., and Presser, S. (1981). Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording, and Context. New York: Academic Press.

Spencer, B.D., Frankel, M.R., Ingels, S.J., Rasinski, K.A., and Tourangeau, R. (1990). NELS:88 Base Year Sample Design Report. Washington, D.C.: National Center for Education Statistics. (NCES 90-463).

Spencer, B.D., and Foran, W. (1991). Sampling probabilities for aggregations, with application to NELS:88 and other educational longitudinal surveys. Journal of Education Statistics, 16(1).

Sudman, S., Bradburn, N.M. (1982). Asking Questions: A Practical Guide to Questionnaire Construction. San Francisco: Jossey-Bass.

Tourangeau, R., and Rasinski, K.A., (1988). Cognitive processes underlying context effects in attitude surveys. Psychological Bulletin, 103, 299-314.

# APPENDIX A

## Documentation of 1990 Instrument Development Meetings

Minutes from 6/4/90 Content Outline Meeting

Agenda from 6/26-27/90 Technical Review Panel Meeting

Memorandum Summarizing 6/13 Meeting in Boulder, CO on Cognitive Test Issues

Memorandum Summarizing the 8/13-14/90 Meeting on Cognitive Test Issues

NORC
6/12/90

# NELS:88 Second Follow-Up Questionnaire Content Outline Meeting

June 4, 1990
Washington, D.C.

Invited Participants:

**Associations**--Rolf Blank (CCSO), Ramsay Selden (CCSO), Sam Husk (Council of Great City Schools), Richard Verdugo (NEA), Bela Rosenberg (AFT), Fred Brigham (NCEA), Joyce McCray (CAPE), Richard Miller (AASA)

**Government Agencies**--Jim Markey (Department of Labor), Richard Marquez (Office of the Secretary of Dropout Advisor), Francine Picoult (Office of Drug Control Policy), Dick Hays (Office of the Secretary Drug Advisor), Bruno Manno (OERI), Milt Goldberg (OR), Oliver Moles (OR), Betty Demarest (OR), Joe Conaty (OR), David Stevenson (OR), Nelson Smith (PIP), Jeanne Griffith (NCES), Paul Planchon (NCES), John Ralph (NCES), Sam Peng (NCES), Dennis Carroll (NCES), Jeremy Finn (NCES), Charles O'Malley (Private Education), Neil Shedd (Office of Indian Education), John Tippeconnic III (Office of Indian Education), Dave Bergeron (OPE), Carmen Simich-Dudegeon (OBEMLA), Jim Stedman (CRS), Audrey Pendleton (OPBE), Elois Scott (OPBE), Bayla White (OMB), Winifred Warnat (OVAE), Mark Schwartz (OVAE), Robert Marshall (OVAE), Jeff Thomas (National Endowment for the Humanities), Thadd Prisco (National Endowment for the Humanities)

**Researchers/Others**--Tony Bryk (University of Chicago), Aaron Pallas (Columbia University), Nancy Karweit (Johns Hopkins University), Ray Valdivieso (Hispanic Policy Development Project), Steven Ingels (NORC), Barbara Schneider (NORC), Leslie Scott (NORC), Fred Newmann (University of Wisconsin), Sue Berryman (Columbia University), Brian Rowan (Michigan State), Ulysses Glee (University of Maryland), Cindy Prince (Maryland State Department of Education), Ivan Charner (National Institute for Work and Learning), Joseph Hawkins (Montgomery County Public Schools)

# NELS:88 Second Follow-Up Questionnaire
# Content Outline Meeting

June 4, 1990
Washington, D.C.

Monday Morning, 4 June 1990

## Welcome and Introduction

The meeting was convened by Chair, Paul Planchon, Director of ESESD, who began with an overview of the National Education Longitudinal Study of 1988 (NELS:88) and its relation to NCES's National Longitudinal Studies program. He informed the audience that NELS:88 represented the third in a series of longitudinal studies that began in 1972 with the National Longitudinal Study of the High School Class of 1972 (NLS-72). This study was followed, in 1980, by the High School and Beyond Sophomore and Senior Cohorts (HS&B). NELS:88, unlike its predecessors, began in 1988 with a substantially younger cohort, a nationally representative sample of eighth grade students. Currently, data collection for the First Follow-Up to NELS:88 (NELS:88 1FU) is being undertaken, and with the recent award of the Second Follow-Up (NELS:88 2FU) contract, work will begin on the 2FU instruments and data collection procedures. Planchon also explained that the purpose of the present meeting, was to engage participants in a discussion of the content areas to be covered by the Second Follow-Up questionnaires. He concluded his presentation by asking all participants to stand and introduce themselves. The introductions began with Deputy Assistant Secretary Bruno Manno, OERI, who bid everyone welcome and a productive meeting.

## Update on the NELS:88 First Follow-Up

Anne Hafner, NELS:88 1FU Project Officer, began the presentations with an update on the 1FU. Participants were told that the study began in the spring of 1989 with a field test of the instruments and data collection procedures. Based on the findings of the field test, the instruments were revised, submitted to OMB, and cleared in December, 1989. Data collection for the Main Study (ongoing at this time) began in January, 1990. Her presentation included a discussion of the final completion rates for district and school contacting, and the current completion rates for the student data collection effort. Hafner also mentioned that imbedded within the First Follow-Up core study of students, teachers and school principals were several supplemental studies of policy-relevant students: the study of LEP/NEP Hispanic students and their parents sponsored by OBEMLA, the study of school effects on student achievement co-sponsored by the MacArthur Foundation and NCES [School Effects Study (SES)], and the study of students who were ineligible for selection into the NELS:88 during the initial base year sampling [Base Year Ineligible Study (BYI)].

## Design Goals and Constraints in the Second Follow-Up Study: Implications for Instrument Development

Steven Ingels of NORC, the NELS:88 Second Follow-Up Project Director, spoke on the implications for instrument development of design and operational considerations. Within the broad study aim of collecting policy-relevant data on educational processes and outcomes, five subsidiary design goals and constraints were identified, that have direct implications for the survey instruments.

1. Limits on questionnaire length, and the need to prioritize as items and themes compete for limited space;

2. The longitudinal character of the primary survey objectives;

3. The time point for the 1992 follow-up -- the completion of secondary schooling -- affords the opportunity both to fulfill the promise of the 1988 baseline measures by gauging the cumulative impact of secondary schooling in the spring of the senior year; and to explore new content areas that establish a broad new baseline at the point of transition to the labor force and post-secondary education, thus permitting measurement of additional outcomes in future waves;

4. The requirement of content stability -- that a core of past items complementary to new items be retained -- not only to provide the continuity that is needed to measure change across waves, but also to sustain comparisons and trend analyses across cohorts, to NLS-72 seniors in 1972 and to HS&B 1980 and 1982 seniors;

5. The need, if data quantity and quality are to be maximized, for articulation between the questionnaires. It is desirable to select from the several potential sources the best respondent or the best records source for the particular data element in question, except where redundancy across questionnaires is desirable to permit analyses of item validities and perceptual differences. The Second Follow-Up includes student/dropout, school administrator, and parent surveys, as well as a teacher survey in at least some subset of schools, and a transcript (and offerings and enrollments data), component as well.

Ingels also touched on issues of confidentiality in public release microdata; and on aspects of the NELS:88 sample design that influence the kinds of research questions the study can address and the kinds of estimates that the NELS:88 database can produce. The Second Follow-Up comprises a nationally representative sample of twelfth grade students in 1992 (with a substantial oversample of Hispanics and Asians), as well as a nationally representative sample of the 1988 eighth grade cohort. However, only in the School Effects Supplement schools will the Second Follow-Up constitute a representative and numerically substantial within-school sample, or a representative school sample. The Second Follow-Up sample design seeks to maximize the effective participation of the 1990 NELS:88 tenth graders in 1992, and participation of several special populations as well. Such special populations will be studied in depth through special questionnaires or supplemental questions: these populations include early graduates and dropouts as well as students added in through sample freshening. A sizable subsample of the ineligible base year students (those excluded because of language barriers or severe physical or mental disabilities) will also be followed.

## New Topics for the NELS:88 Second Follow-Up: Overview

Following Steven Ingels was Jeanne Griffith of NCES, who presented an overview of new 2FU questionnaire content areas, as well as content areas covered in previous NELS:88 instruments--both the base year and 1FU.

The potential content areas to be covered in the Second Follow-Up instruments -- the issues for discussion at today's meeting -- Griffith explained, included the somewhat indeterminate construct of restructuring: school organization/climate; access and choice in postsecondary education and the workplace; equity issues, such as consequences of tracking, allocation of school resources, and family and community support; student engagement; and teacher professionalism and school/classroom experiences.

## Access and Choice in Postsecondary Education and the Workplace

The next speaker to address the group was C. Dennis Carroll, NCES Postsecondary Longitudinal Studies Branch Chief. Carroll's presentation centered on the need for developing questionnaire items that would gather data on students' access to and choice of postsecondary education/schools. Defining access and choice simply as whether or not young people go to college -- how and when they decide what colleges to go to -- Carroll recommended that the following topics be addressed in the student questionnaire:

- Decision and timing -- measures of the postsecondary decision-making process and timing of the decision to go to college;
- Early intervention of student financial assistance -- measures of whether students who said during the First Follow-Up that they would be attending college but have changed their mind by the Second Follow-Up, as well as students who said they would not be attending college but who by the twelfth grade say that they now will (i.e., "late bloomers");
- Parental aspirations -- this would include, perhaps, measures of parental involvement, if any, in influencing whether or not to go to college;
- Preparation -- traditional behavioral measures of preparation, such as what courses students take, as well as less conventional behavioral measures, such as students' requests for college catalogs and number of campus visits;
- Admission tests -- seen as another behavioral measure of "preparatoriness," -- this topic includes measures of whether the student has taken or is preparing to take standard admissions tests (ACT, SAT), as well as specific state admissions tests; actual test scores; and the names of the institutions to which the student has sent his/her scores;
- Application -- measures of the timing of students' application, e.g., early admission; the number of applications submitted; names of the first three schools of choice (this measure may be used for assessing "flyers" or schools that students would like to attend but

could never get into); and schools that accepted the student;

- Student financial aid -- measures of whether the student was offered financial aid at the school of acceptance; type of financial aid offered (e.g., grants, loans, work study); and timing of financial aid (e.g., at time of acceptance or later);

- Campus visits -- measures of the number and timing of campus visits; number of campus interviews.

For the dropout questionnaire, Carroll recommended inclusion of items that will look more closely at proprietary institutions or vocational/technical schools. Two specific areas of interest were:

- Proprietary recruitment -- measures of how an institution recruited the student, especially--what abilities/skills did the institution state the student must possess for acceptance into the program (e.g., a certain score on an admissions test)?;

- Remediation -- measures of if and when the student was identified as needing remedial services.

David Stevenson, Office of Research, opened the discussion by asking if the Second Follow-Up questionnaires would contain any items on students' and/or dropouts' transition into the labor force. Stevenson acknowledged that, as in HS&B, the early administration of the senior year questionnaire -- as early as February 1992, in a data collection period that runs from early February to late May and peaks in March -- inhibits the collection of accurate data on certain issues of importance, such as whether the student/dropout has thought about work after high school, and even, where the student will attend college and the kind, if any, of financial aid offered. Continuing, Stevenson, stated that the collection of this kind of data, two years later, would be adversely affected by students' recall, and he recommended that the same problematic design of NLS-72 and HS&B not be repeated for the Second Follow-Up. It might be best, he suggested, to administer a subset of items to a subsample of students in the October following their senior year.

Fred Newmann, University of Wisconsin, noted the implicit societal notion that the only alternative after high school is college, and that this notion tends to carry over to research. Dignity, he stated, does not only lie with college acceptance but also with work; the group should not lose sight of the importance of assessing this issue. Agreeing, Sue Berryman, Columbia University, suggested that items be developed that measure how schools are organized to support the transition from high school to work. Joyce McCray, CAPE, returning to the issue of the methodological problem inherent in measuring students' labor knowledge/behaviors, stated that students in the middle academically will not know what they will be doing after high school. And, she supported Stevenson's call for two data collection periods. Tony Bryk, University of Chicago, suggested that if the measurement of students' transition to the workplace is important, then only those items that could be supported by the present design should be kept or developed.

With this in mind, Stevenson suggested the inclusion of items which capture students' understanding of the labor force/market, specifically, questions on whether students feel they are learning material that will be applicable to the jobs they seek. Stevenson noted that some labor force items may be gleaned from the Department of Labor's NLS/Y. Berryman reminded the group that the student work items represent only one half of the picture. What schools know, and hence, train students about the world of work constitutes the other half of the picture, and also needs to be inquired into by NELS:88.

In concluding, Fred Newmann attempted to summarize the group's concerns and recommendations: 1) to look more broadly at students' perspective of the future; include items on both college and work; 2) to include items on students' perceptions of the labor market: what opportunities are out there; what credentials are needed; 3) to include items on students' present and past work experiences; and 4) to include items on the school's and parents' role in providing a support system with respect to students' transition to the work place. Newmann also noted the need for addressing these topics in the student, parent and school administrator questionnaires. Stevenson suggested again that serious consideration be given to collecting data at two points in time. Barbara Schneider, NORC, noted that the methodological limitations of the present design could be overcome without changing the design. For example, rather than adding a second formal data collection period in October, those labor and college items most susceptible to timing could be asked of a subsample of respondents during a scheduled address update.

## Equity Issues: Consequences of Tracking, Allocation of School Resources, and Family and Community Support

The next speaker was Nancy Karweit of Johns Hopkins University who spoke on the need for several new content areas, and possible design modifications, to assess the perennial issue of equity. These content areas, potential items, and design concerns were:

- Practices and effects related to restructuring, specifically with respect to equity and assisting disadvantaged students -- Restructuring may be defined as any change in governance, management and administration in order to facilitate a more responsive learning environment. New areas include measures of alternatives to tracking, and ability grouping that schools may have implemented. It is necessary to measure both what the alternatives are and the results of adopted alternatives or restructuring. One design concern is the need for using multiple respondents for capturing change and the results of change. For example, if you ask principals whether a certain administrative change has been adopted they will say yes. However, if you were to ask the same question of teachers, they would respond no. In essence, one needs to ask principals what they think is happening, ask teachers what they are actually doing, and ask students how they perceive any changes. Only with such triangulation can the construct of restructuring and the results of restructuring be measured.

- Continued isolation of students at risk from knowledge of the future -- This area concerns the questions of how students get access to knowledge about other worlds and what kind of knowledge they have about other worlds other than their own.

- Access to support systems and resources -- More items need to be developed for tapping: who the student turns to during times of trouble; the existence of integrated support systems, such as those between the school and community; and whether students have access to summer school or tutors or counselors. The issue of access to resources touches upon such broad topics as teacher training, and restructuring practices, such as reduced class size. Here, again items need to be developed that address access issues in both the teacher and student questionnaires.

- Changes in the incentive structure -- This issue concerns the measurement of students' perceptions of incentives and rewards. Specifically, what are students' rewards for working hard and how do students perceive such incentives and rewards -- are they perceived as fair or unfair?

- Parental network -- This issue concerns the extent to which the parents know each other, the schools' policies, and the extent to which the school knows parents' concerns.

Barbara Schneider initiated the ensuing discussion by suggesting that items be added which measure students' awareness of the consequences of taking or not taking certain courses. Fred Newmann recommended the inclusion of items that tap the school administrator's perception of tracking and grade weighting. Jeremy Finn, AASA Fellow, expressed concern over the exclusionary tone of the student questionnaire. The student questionnaire, he noted, is geared toward students who have given such issues as course taking behavior serious consideration when in reality a sizable portion of students probably have not. He advocated exploration of the "don't know" categories for some items through a series of follow-up questions, such as "Why don't you know;" "Do you have a counselor?"

Nancy Karweit instanced another issue, vocational education and schools. Karweit suggested adding items that measure how students learn different vocations, such as bricklaying. Participants agreed, and it was further suggested that students be asked about their vocational classes and experiences, and perhaps, even add a vocational teacher supplement. Fred Newmann commented on the other side of vocational course taking -- the establishment of the work ethic. Other vocational education issues discussed included the measurement of different types of vocational programs and schools, for example, pull out programs, vocational programs within traditional schools and free standing vocational schools; how the student got into vocational school --is he/she attending through a magnet or choice program to which the student applied or is it the only school in the area.

## Students' Engagement

Beginning with a definition of student engagement -- the psychological investment in acquiring knowledge and mastering skills -- Fred Newmann, University of Wisconsin, discussed his theoretical research on the construct of student engagement and potential ways for measuring it. The construct, Newmann continued, is composed of two dimensions: a psychological component which may be further defined as students' sense of intrinsic interest in school and a behavioral component which is the actual work and effort that students put forth to master school material. The conditions which are necessary for engagement to occur are a sense of social membership in the organization and challenging work. The concept of social membership encompasses the issue of social support -- the availability and use of mentors, tutors and counselors-- but it is best understood as a bipolar cohesiveness-alienation dimension. Operationally, it concerns the extent to which students see themselves as a member of the school social life or simply life. The second condition necessary for engagement to occur, challenging work, has four aspects: the extent to which students feel challenged to use their mind; an active-passive aspect -- are students actively participating in learning either through discussion or hands-on experience; a depth versus breadth aspect -- are students given time to learn a subject in depth or do they skip superficially over a variety of topics; and the extent of substantive classroom conversation or the verbal or written exchange of ideas. Newmann suggested that new items be developed for assessing whether these conditions exist, and even, perhaps, for measuring the conditions that produce a high sense of social membership and challenging work. He also suggested the use of achievement on authentic tasks and inauthentic tasks as potential outcome measures.

In the discussion that followed, Jeremy Finn brought up the idea of using behavioral measures of engagement, such as participation in clubs. Newmann agreed, but also stated the need to get at the "quality of participation." It is not enough, he stated, to just measure how many times a student raised his/her hand; measures must also be taken of what was said. Tony Bryk noted that the assessment of an engaging environment may be best addressed as a school climate factor. Barbara Schneider noted the difficulty of measuring student engagement in twelfth grade particularly since the end of twelfth grade marks the beginning of disengagement from school. Many schools, she noted, do not even require seniors to take final exams.

## Teacher Professionalism and School/Classroom Experiences

Brian Rowan of Michigan State University spoke next, on the topics of teachers and teaching and how these issues relate to classroom experiences and outcomes. These topics are fundamentally important, he explained, because teachers represent the midpoint between schools and students; school structure affects teachers, who in turn, affect students. Rowan described three areas of interest:

1. Instructional practices/teaching strategies -- This area concerns the measurement of a set of teaching strategies that are related to student success at higher order learning or problem solving. This set of successful teaching strategies may be called "teaching for understanding." Teaching for understanding may be examined through items on either the teacher or student questionnaire that inquire about such things as the content areas covered in a particular class; how teachers prepare for class; homework assigned; how teachers assess knowledge (i.e., assign grades); and the

openness/closeness of the student-teacher relationship;

2. How engaged teachers are in their work -- This area concerns not only what teachers are doing inside of work or the classroom but also what they are engaged in outside of work. Teaching for understanding may involve more teacher preparation outside of class, and/or a specific academic background;

3. The relationships between teachers and students -- To engage in teaching for understanding, teachers need to develop more personal relationships with their students. Teachers need to talk and associate with students on several levels in order to get them to "open up".

Rowan also spoke on the effect of school contextual factors on teachers' work and the notion of teaching for understanding. For example, knowledge of students' track assignment can affect teachers' approach to teaching. Another factor is the assignment of teachers to classes for which they have no to minimal experience teaching. Rowan also discussed interest in identifying a set of school factors that may be directly related to teachers' ability to teach for understanding, such as supportive leadership, collaborative relationships among teachers, teacher control over instructional strategies, and high standards to which everyone is committed.

Group discussion began with Barbara Schneider who noted the need for operationalizing such difficult process-oriented constructs as "teaching for understanding" and Newmann's construct of student engagement in ways that would relate to the important activities and events that seniors experience, such as taking the ACT, asking a particular teacher for a recommendation to college. Tony Bryk mentioned that since the design of the Second Follow-Up does not allow for a random sample of teachers within schools, or even, necessarily, a teacher survey at all, except within School Effects Supplement (SES) schools, the assessment of teacher effects is a moot point. He suggested the design modification of drawing a random sample of teachers with the SES schools and administering only teacher effect items.

## School Organization/Climate--Restructuring

The final speaker, Tony Bryk of the University of Chicago, spoke on restructuring. Bryk sketched the topics currently being addressed by researchers interested in restructuring. At the national level, restructuring issues concern measurable results, such as the achievement of educational goals. At the community and school level, issues of restructuring primarily concern the coordination of resources among social/business agencies. At the level of the school, the concept encompasses such issues as how the school interacts with the community; strengthening of the social ties between parents and key actors within the schools (e.g., parent involvement/volunteerism). Inside the school, restructuring efforts focus on issues of student engagement; teacher professionalism and development; and teacher control over school governance issues and leadership. And, at the lowest level, restructuring issues concern changes that occur in the classroom. Implications, Bryk continued, for item development include assessment of: the activities that schools engage in; who initiated the activity (e.g., district, state); teachers' and students' perception of change; and actual change.

The group's discussion of restructuring centered primarily on the technical problems inherent in the measurement of the concept. Fred Newmann pointed out that many teachers and school administrators are unaware of policy or structural changes that occur within their schools. One way to get around this problem may be to develop a profile of an ideal, restructured school, and then, assess the extent to which schools deviate from this ideal. The group also discussed additional restructuring issues that they would like to see covered by the instruments, such as the effect of teacher morale on student achievement and engagement, and school administrators' sense of the school's purpose.

## Session Summary

In conclusion, Barbara Schneider highlighted the major content areas of interest and potential design changes: a follow-up subsample of students to ascertain employment or enrollment in postsecondary school; NAEP-NELS test score equating; random sampling of SES teachers for assessing teacher and school effects; the administration of several items on restructuring to state and district officials at the time of state and district contacting; the inclusion of items for measuring support systems, natural variation, the quality of school experiences, and teaching for understanding.

Leslie Scott, NORC,
Recorder.

NELS 88
National
Education
Longitudinal
Study of 1988

Sponsored by The National Center For Education Statistics
United States Department of Education
Conducted by NORC
A Social Science Research Center

## NELS:88 Second Follow-Up
## Technical Review Panel (TRP) Meeting
June 26-27, 1990

Room 326, 555 New Jersey Ave., Washington, D.C.

---

**Tuesday, June 26, 1990**

| | |
|---|---|
| 10:00-10:15 | Welcome and Introduction -- Paul Planchon, NCES |
| 10:15-10:30 | Update on Base Year (BY) Products: NORC and MPR Reports -- Jeff Owings, NCES |
| 10:30-11:15 | Update First Follow-Up Progress; Overview of the Second Follow-Up: Components and Milestones; 2FU Field Test Design -- Steven Ingels, NORC |
| 11:15-12:15 | Cognitive Test Battery: Lessons from BY and 1FU; Issues for 2FU -- Don Rock, ETS; Commentator: Lyle Jones, L.L. Thurstone Psychometric Lab, University of North Carolina |

**12:15 Working Lunch: Discussion on Testing Issues**

| | |
|---|---|
| 1:00-3:00 | Student Questionnaire: Themes, Issues, Items -- Jerry Bachman, University of Michigan, ISR, and Barbara Schneider, NORC, Instrumentation |
| 3:00-3:15 | Break |
| 3:15-4:30 | Parent Questionnaire: Themes, Issues, Items -- Nancy Karweit, Johns Hopkins University, and Barbara Schneider, NORC |

**Wednesday, June 27, 1990**

| | |
|---|---|
| 8:30-9:00 | Coffee, Juice and Pastries |
| 9:00-9:30 | Supplements -- Iris Rotberg, NSF; Carmen Simich-Dudgeon, OBEMLA |
| 9:30-10:45 | Base Year Descriptive Summary: News Conference -- Anne Hafner, NCES |
| 10:45-11:00 | Break |
| 11:00-12:00 | School Questionnaire: Themes, Issues, Items -- Tony Bryk, University of Chicago, and Barbara Schneider, NORC |
| **12:00 Working Lunch:** | **School Questionnaire, continued** |
| 1:30-3:30 | Dropout Questionnaire: Themes, Issues, Items -- Aaron Pallas, Columbia University, and Barbara Schneider, NORC |
| 3:00-3:15 | Summary and Wrap-Up -- Steven Ingels, NORC |

July 6, 1990

TO: Addressees

FROM: Anne Hafner
Project Officer, NELS:88

RE: June 13th NELS:88 Boulder meeting

The purpose of this memo is to summarize discussions re: NELS:88 cognitive tests and topic coverage held at a meeting in Boulder, Colorado on June 13th. A writeup of a subsequent meeting held at the NELS:88 Technical Review Panel in Washington on June 26th is forthcoming. Participants at the June 13th meeting included: Steven Ingels, NORC; Don Rock, ETS; Judy Pollack, ETS; Anne Hafner, NCES; Steven Gorman, NCES; Ron Anderson, University of Colorado; Leigh Burstein, UCLA; Tej Pandey of the California State Department of Education; and Gordon Ensign of the Washington State Dept. of Education.

## BOULDER MEETING

### Test and Item Specifications

Rock presented an overview of the historical background of NCES' use of cognitive tests within the longitudinal studies program. He discussed various constraints placed upon test development, including limiting floor and ceiling effects, the necessity of including some HS&B items so as to be able to equate NELS:88 and HS&B, and that the tests should be unspeeded. Don distributed "Test Specifications" which described specs for tests, along with the math framework (process x content grids) for NELS:88 grades 8 and 10, along with NAEP's framework. The reading, science and history/government frameworks were also presented.

The content categories for the math test were discussed, as they related to the NAEP subscales. There was a brief discussion of item types, including the extensive use of quantitative comparison items in HS&B and NELS:88. Rock explained that since HS&B contained only quantitative comparison items, it was necessary to include quite a few of these items on NELS:88, in order to equate the two. These QC items were drawn from the SAT, and there might be a problem with familiarity with format with 8th graders. Leigh Burstein suggested using more IEA math items in the second followup, as IEA-math test has many good items. Rock plans to add pre-calculaus and function items to the 12th grade test.

It was suggested that some type of spiraling of the tests be done in order to add more math items, especially NAEP items. For example, one half of the sample could take the social studies or reading test and the other half would not take the social studies or reading test but would take a longer math test which included NAEP items.

## Science Test

Ron Anderson from the University of Colorado discussed some problems with science as it is currently taught and assessed. Most multiple choice test items do not get to the understanding level. He noted that levels generally used include knowing science, conducting inquiries and solving problems. He suggested adding physics quex to the 12th grade test and using fewer earth and life science items. The content areas discussed were life science, earth science, and physical science/chemistry. Suggested coverage topics included chemistry, physics, and general science.

## Proficiency Scales

The math and reading proficiency scales used for the base year were discussed. The same scales will probably be used for first followup. For second followup, it is likely that a new higher level which may encompass pre-calculus and function items will be added. Tej Pandey brought up the fact that some high school math classes teach "integrated mathematics" in which geometry and algebra are taught together in an integrated fashion. Rock explained that the proficiency scales incorporated both content and process.

## Open Ended Items

There was a dissussion of the use of open ended math and science items in the 12th grade test battery. There is some money in the second followup contract for developing and field testing free response items. National Science Foundation may be willing to fund administration and scoring in full scale NELS in 1992. Adding two open ended items would require 15-20 minutes per child. Various ways of freeing up 15-20 minutes were discussed, including pullout groups which will be timed separately, and dropping the reading test in the 12th grade (20 minutes). The arguments for dropping the reading test were the following: a) reading scores don't change much from 10th to 12th grade; b) reading is an aptitude test which is meant to be used as a control variable; c) math and science are more important to measure than other subjects in 12th grade; d) it is important to measure performance on open-ended items, thus space should be freed up and e) most states test reading in 12th grade.

There was general agreement that any open ended items should be separately timed. Tej Pandey said that NCES could probably use some of California's open ended items, as well as its scoring protocols.

One idea that was discussed was that of doing the pullout only in SES schools (250 schools).

NAEP/NELS:88 Equating

There was a discussion of equating NAEP and NELS:88, as well as different ways of doing this. One suggestion was to take a sample of the NELS:88 kids, and have one half of the kids take the NELS:88 items and one random half a NAEP booklet. Then the two could be put on an equipercentile scale. Rock seemed to think we could equate the two by giving NELS:88 kids more NAEP items, but there was not general agreement on this point.

Topic Coverage/Other Topics

Leigh Burstein discussed various ways of measuring topic coverage. One way suggested was to measure the number of periods spent on a topic: None, 1 or 2, one week, 1 to 3 weeks, 4 weeks or more. It was felt that a measure which got to the time spent on a topic would be more useful. Burstein suggested that perhaps we should try to survey all science and math teachers in the students' school. Another suggestion was to survey teachers of students' 11th and 12th grade math and science classes, as many students will not be taking math or science in 12th grade.

It was suggested that NCES look at Iris Weiss' teacher questionnaire for items, and talk to the Chiefs and ask them for possible items.

cc: Emerson Elliott
Paul Planchon
Jeff Owings
Steve Ingels

Addressees: Ron Anderson
Leigh Burstein
Gordon Ensign
Steve Gorman
Tej Pandey
Gary Phillips
Judy Pollack
Don Rock

NORC

A Social Science Research Center
University of Chicago

1155 East 60th Street, Chicago, IL 60637
312/702-1200 Fax 312/702-0857

**TO:** Members of the NELS:88 Technical Review Panel; and: Leigh Burstein, Tom Romberg, Tej Pandey, Dick Berry, Senta Raizen

**DATE:** October 16, 1990

**FROM:** Steven Ingels, 312/702-8999
NELS:88 Project Director

**SUBJECT:** Psychometric Plans for the 1991 Field Test, NELS:88 Second Follow-Up

cc: Jeff Owings, Peggy Quinn, Anne Hafner, Gary Phillips, Steven Gorman, NCES; Donald Rock, Judith Pollack, ETS

Thank you for your past contributions to our planning for the NELS:88 Second Follow-Up. We want to keep you closely informed of the latest developments in our plans for the field test, and give you an opportunity to react to these plans. Below, we sketch the psychometric background to the spring 1991 field test, the basic assumptions about the goals and logistics of the field test, and the basic plan for implementing the field test design. If you have particular suggestions or comments, please call in or fax them to us as soon as possible, and by October 24, 1990, at the very latest. Your comments can be directed either to me or to Peggy Quinn at NCES (202) 357-6743.

Background.

The following conclusions were reached at the August 13-14 meetings at NCES on cognitive testing issues for NELS:88.

(1) The study should seriously explore in the field test the possibility of allocating testing time in the main study (1992 twelfth grade administration) to free response items in both mathematics and science. Such items should be field-tested and their potential for contributing to the main study cognitive test battery thoroughly weighed in the field test report that will be considered when the Technical Review Panel is convened in June of 1991.

(2) If after reviewing the field test results, it was decided that free response items be added to the cognitive test battery in the main study, extra time necessary for the administration of free response items could be secured by not repeating the reading comprehension test. There is a strong two-fold rationale for omitting reading: (a) in most analyses, reading would primarily be used as a control variable; (b) only minor gains could be expected in reading ability between grades ten and twelve. (It should be noted however that over 3,000 First Follow-Up sample members and an additional 1200 students new to the Second Follow-Up [via sample freshening] will

lack a 1990 reading score. Given the critical need to measure reading comprehension, it may well be necessary to administer a reading test to this population, even if the reading test is dropped for twelfth graders for whom a 1990 reading score is available).

(3) It would also be highly desirable to equate NAEP and NELS:88 twelfth grade mathematics scores. This goal should be pursued if at all possible (that is, if funding permits). For this reason, the 1991 field test should include additional items (especially geometry items, which were not well represented in the NELS:88 1990 First Follow-Up cognitive tests) that will bring NELS:88 and NAEP mathematics test specifications as close as possible.

Guiding Assumptions.

- The maximum time available for cognitive testing will be 90 minutes.
- The field test sample will yield approximately 3,000 test observations from one hundred field test schools in the five field test states (California, New York, Texas, Florida, and Illinois).
  - One thousand sample members will be drawn from the longitudinal sample of past field test participants;
  - Freshened (out of modal sequence) students and an augmented sample of twelfth graders will provide an additional two thousand sample members for the 1991 field test;
- In 1991, given the large number of non-longitudinal sample members, and the radical differences between the 1991 tests and those pretested in 1987 and 1989, randomization of assignment to test forms within schools will be the paramount aim, and will supplant the longitudinal analyses that informed the 1989 pretest;
- A major purpose of the 1991 pretest will be to investigate both the operational feasibility and the "value added" of incorporating open-ended science and mathematics items into the 1992 main study test administration.

Field Test Plans.

It is suggested that a spiralled design be used that incorporates both previously administered multiple choice items (mci's), "new" multiple choice items, and constructed response items (cri's). It is further suggested that the cri's be restricted to the mathematics and science content areas. In the mathematics and science tests, additional time will be allocated to the task of carrying out pacing instructions. (Experience with the NAEP constructed response format has suggested that students may not--in the absence of pacing--allocate their time wisely when faced with open-ended formats). Further time will be reserved so that students can respond to questions that will determine whether (a) the student sufficiently understood the task that was being posed and (b) whether the student felt more (or less) comfortable with the open-ended assessment technique than with traditional multiple choice assessment items.

Potential field test spirals might be designed as follows:

| | Section 1:<br>25 minutes | Section 2:<br>17 minutes | Section 3:<br>4 12-minute items = 48 min. |
|---|---|---|---|
| Form I (Math A) | 12 old m.c. items, plus about 15 new mostly geometry (some NAEP 1991) | 15 new m.c. items, more difficult topics (some NAEP 1992) | 4 free response items separately timed, including motivation/ difficulty questions |
| Form II (Math B) | same items as Form I | same items as Form I | 4 more free response math items |
| Form III (Sci/Hist A) | 10 old m.c. science items, plus about 15 new, more difficult | 10 old m.c. hist/cit/geog plus about 12-15 new, more difficult | 4 free response science items, separately timed plus questions |
| Form IV (Sci/Hist B) | same items as Form III | same items as Form III | 4 more free response science items |
| Form V (Reading) | 2 new (harder) passages with m.c. items | 2 old reading passages with m.c. items | 4 free response reading items, separately timed plus questions |

## APPENDIX B

### Documentation in Support of
### Student Questionnaire Experiment Analyses

```
Crosstabulation:      Q8A       # TIMES R WAS LATE FOR SCHOOL
                   By RTYPE     response type

 +------------------------+---------------+---------------+---------------+
 |           Count        | open          | closed        |               |
 |   RTYPE-> Col Pct      |               |               | Row           |
 |           Residual     |           .00 |          1.00 | Total         |
 +------------------------+---------------+---------------+---------------+
 |    Q8A                 | 112           | 86            | 198           |
 |                    .00 | 23.1%         | 17.5%         | 20.3%         |
 |      NEVER             | 13.7          | -13.7         |               |
 +------------------------+---------------+---------------+---------------+
 |                   1.00 | 130           | 189           | 319           |
 |      1-2 TIMES         | 26.9%         | 38.5%         | 32.7%         |
 |                        | -28.4         | 28.4          |               |
 +------------------------+---------------+---------------+---------------+
 |                   2.00 | 153           | 124           | 277           |
 |      3-6 TIMES         | 31.6%         | 25.3%         | 28.4%         |
 |                        | 15.5          | -15.5         |               |
 +------------------------+---------------+---------------+---------------+
 |                   3.00 | 21            | 36            | 57            |
 |      7-9 TIMES         | 4.3%          | 7.3%          | 5.8%          |
 |                        | -7.3          | 7.3           |               |
 +------------------------+---------------+---------------+---------------+
 |                   4.00 | 68            | 56            | 124           |
 |      OVER 10 TIMES     | 14.0%         | 11.4%         | 12.7%         |
 |                        | 6.4           | -6.4          |               |
 +------------------------+---------------+---------------+---------------+
 |                 Column | 484           | 491           | 975           |
 |                  Total | 49.6%         | 50.4%         | 100.0%        |
 +------------------------+---------------+---------------+---------------+

  Chi-Square    D.F.     Significance      Min E.F.     Cells with E.F.< 5
  ----------    ----     ------------      --------     ------------------
   22.42203      4           .0002          28.295        None

Number of Missing Observations =      21
```

```
Crosstabulation:     Q8D        # TIMES R WAS IN TROUBL FOR IGNORING RUL
                  By RTYPE      response type

+------------------------+-------------+-------------+-----------+
|           Count        | open        | closed      |           |
|  RTYPE->  Col Pct      |             |             | Row       |
|           Residual     |         .00 |        1.00 | Total     |
+------------------------+-------------+-------------+-----------+
|  Q8D                   | 341         | 310         | 651       |
|                    .00 | 70.2%       | 62.6%       | 66.4%     |
|    NEVER               | 18.5        | -18.5       |           |
+------------------------+-------------+-------------+-----------+
|                   1.00 | 95          | 130         | 225       |
|    1-2 TIMES           | 19.5%       | 26.3%       | 22.9%     |
|                        | -16.5       | 16.5        |           |
+------------------------+-------------+-------------+-----------+
|                   2.00 | 40          | 36          | 76        |
|    3-6 TIMES           | 8.2%        | 7.3%        | 7.7%      |
|                        | 2.3         | -2.3        |           |
+------------------------+-------------+-------------+-----------+
|                   3.00 | 3           | 7           | 10        |
|    7-9 TIMES           | .6%         | 1.4%        | 1.0%      |
|                        | -2.0        | 2.0         |           |
+------------------------+-------------+-------------+-----------+
|                   4.00 | 7           | 12          | 19        |
|    OVER 10 TIMES       | 1.4%        | 2.4%        | 1.9%      |
|                        | -2.4        | 2.4         |           |
+------------------------+-------------+-------------+-----------+
|                 Column | 486         | 495         | 981       |
|                  Total | 49.5%       | 50.5%       | 100.0%    |
+------------------------+-------------+-------------+-----------+

Chi-Square    D.F.     Significance      Min E.F.     Cells with E.F.< 5
----------    ----     ------------      --------     ------------------
   9.96522       4            .0410         4.954     1 OF    10 ( 10.0%)

Number of Missing Observations =        15
```

Crosstabulation: Q10 # OF DAYS MISSED DURING LAST ABSENCE
By RTYPE response type

| RTYPE-> Count<br>Col Pct<br>Residual | open<br>.00 | closed<br>1.00 | Row<br>Total |
|---|---|---|---|
| Q10<br>.00<br>NEVER | 23<br>9.0%<br>11.5 | 0<br>.0%<br>-11.5 | 23<br>4.5% |
| 1.00<br>1-2 DAYS | 201<br>78.5%<br>-11.8 | 223<br>87.8%<br>11.8 | 424<br>83.1% |
| 2.00<br>3-4 DAYS | 17<br>6.6%<br>-.1 | 17<br>6.7%<br>.1 | 34<br>6.7% |
| 3.00<br>5-10 DAYS | 10<br>3.9%<br>.5 | 9<br>3.5%<br>-.5 | 19<br>3.7% |
| 4.00<br>11-15 DAYS | 4<br>1.6%<br>1.0 | 2<br>.8%<br>-1.0 | 6<br>1.2% |
| 5.00<br>16-20 DAYS | 1<br>.4%<br>-.0 | 1<br>.4%<br>.0 | 2<br>4% |
| 6.00<br>21 OR MORE DAYS | 0<br>.0%<br>-1.0 | 2<br>.8%<br>1.0 | 2<br>.4% |
| Column<br>Total | 256<br>50.2% | 254<br>49.8% | 510<br>100.0% |

```
Chi-Square      D.F.      Significance      Min E.F.      Cells with E.F.< 5
----------      ----      ------------      --------      ------------------
  26.85338        6           .0002             .996      6 OF    14 ( 42.9%)

Number of Missing Observations =     486
```

```
Crosstabulation:     Q24A1     TOTAL TIME SPENT ON HW IN SCHOOL
                  By RTYPE     response type

              Count     open          closed
   RTYPE->    Col Pct                               Row
              Residual        .00           1.00    Total
   Q24A1                       54             41       95
                  1.00      11.2%           8.5%     9.8%
     NONE                     6.5           -6.5

                  2.00        133            197      330
     1 HR OR LESS           27.5%          40.6%    34.1%
                            -31.8           31.8

                  3.00        115            123      238
     2-3 HOURS              23.8%          25.4%    24.6%
                             -3.9            3.9

                  4.00         91             71      162
     4-6 HOURS              18.8%          14.6%    16.7%
                             10.1          -10.1

                  5.00         23             19       42
     7-9 HOURS               4.8%           3.9%     4.3%
                              2.0           -2.0

                  6.00         24             15       39
     10-12 HOURS             5.0%           3.1%     4.0%
                              4.5           -4.5

                  7.00         14             10       24
     13-15 HOURS             2.9%           2.1%     2.5%
                              2.0           -2.0

                  8.00         30              9       39
     OVER 15 HOURS           6.2%           1.9%     4.0%
                             10.5          -10.5

                Column        484            485      969
                 Total      49.9%          50.1%   100.0%

 Chi-Square     D.F.     Significance       Min E.F.     Cells with E.F.< 5
 ----------     ----     ------------       --------     ------------------
   31.36035       7          .0001           11.988        None

Number of Missing Observations =       27
```

```
Crosstabulation:      Q24A2      TOT TIME SPENT ON HW OUT OF SCHL
                   By RTYPE      response type

            Count     open          closed
 RTYPE->    Col Pct                                Row
            Residual          .00            1.00  Total
Q24A2                          27              40      67
                  1.00       5.6%            8.2%    6.9%
  NONE                       -6.4             6.4

                  2.00         93              95     188
  1 HR OR LESS              19.1%           19.4%   19.3%
                               -.7              .7

                  3.00        124             124     248
  2-3 HOURS                 25.5%           25.4%   25.4%
                               .4              -.4

                  4.00        108              83     191
  4-6 HOURS                 22.2%           17.0%   19.6%
                             12.8           -12.8

                  5.00         34              48      82
  7-9 HOURS                  7.0%            9.8%    8.4%
                             -6.9             6.9

                  6.00         43              44      87
  10-12 HOURS                8.8%            9.0%    8.9%
                              -.4              .4

                  7.00         18              30      48
  13-15 HOURS                3.7%            6.1%    4.9%
                             -5.9             5.9

                  8.00         39              25      64
  OVER 15 HOURS              8.0%            5.1%    6.6%
                              7.1            -7.1

                Column        486             489     975
                 Total      49.8%           50.2%  100.0%

 Chi-Square    D.F.     Significance       Min E.F.     Cells with E.F.< 5
 ----------    ----     ------------       --------     ------------------
   14.27106       7          .0466          23.926         None

Number of Missing Observations =      21
```

```
Crosstabulation:      Q24C2     TIME SPENT ON SCIENCE HW OUT OF SCHOOL
                    By RTYPE     response type

               Count   | open          | closed        |
  RTYPE->      Col Pct |               |               | Row
               Residual|           .00 |          1.00 | Total
  ---------------------+---------------+---------------+--------
  Q24C2                |       64      |      139      |   203
                  1.00 |       23.4%   |      34.2%    |  29.9%
    NONE               |      -17.8    |      17.8     |
  ---------------------+---------------+---------------+--------
                  2.00 |      109      |      128      |   237
    1 HR OR LESS       |       39.8%   |      31.5%    |  34.9%
                       |       13.5    |     -13.5     |
  ---------------------+---------------+---------------+--------
                  3.00 |       65      |       83      |   148
    2-3 HOURS          |       23.7%   |      20.4%    |  21.8%
                       |        5.4    |      -5.4     |
  ---------------------+---------------+---------------+--------
                  4.00 |       30      |       36      |    66
    4-6 HOURS          |       10.9%   |       8.9%    |   9.7%
                       |        3.4    |      -3.4     |
  ---------------------+---------------+---------------+--------
                  5.00 |        1      |       13      |    14
    7-9 HOURS          |         .4%   |       3.2%    |   2.1%
                       |       -4.6    |       4.6     |
  ---------------------+---------------+---------------+--------
                  6.00 |        3      |        6      |     9
    10-12 HOURS        |        1.1%   |       1.5%    |   1.3%
                       |        -.6    |        .6     |
  ---------------------+---------------+---------------+--------
                  7.00 |        1      |        0      |     1
    13-15 HOURS        |         .4%   |        .0%    |    .1%
                       |         .6    |       -.6     |
  ---------------------+---------------+---------------+--------
                  8.00 |        1      |        1      |     2
    OVER 15 HOURS      |         .4%   |        .2%    |    .3%
                       |         .2    |       -.2     |
  ---------------------+---------------+---------------+--------
                Column |      274      |      406      |   680
                 Total |       40.3%   |      59.7%    | 100.0%

  Chi-Square    D.F.     Significance       Min E.F.      Cells with E.F.< 5
  ----------    ----     ------------       --------      ------------------
    19.35887      7          .0071              .403      5 OF    16 ( 31.3%)

Number of Missing Observations =      316
```

```
Crosstabulation:     Q24E1     TIME SPENT HIST/SOC STUD HW IN SCHOOL
                  By RTYPE     response type

+------------------------+-------------+-------------+----------+
|              Count     | open        | closed      |          |
|   RTYPE->    Col Pct   |             |             | Row      |
|              Residual  |         .00 |        1.00 | Total    |
+------------------------+-------------+-------------+----------+
|  Q24E1                 |      118    |       40    |   252    |
|                   1.00 |      32.8%  |       8.2%  |   31.5%  |
|    NONE                |      4.6    |      -4.6   |          |
+------------------------+-------------+-------------+----------+
|                   2.00 |      140    |      201    |   341    |
|    1 HR OR LESS        |      38.9%  |      45.7%  |   42.6%  |
|                        |     -13.4   |      13.4   |          |
+------------------------+-------------+-------------+----------+
|                   3.00 |       42    |       73    |   115    |
|    2-3 HOURS           |      11.7%  |      16.6%  |   14.4%  |
|                        |      -9.8   |       9.8   |          |
+------------------------+-------------+-------------+----------+
|                   4.00 |       43    |       23    |    66    |
|    4-6 HOURS           |      11.9%  |       5.2%  |    8.3%  |
|                        |      13.3   |     -13.3   |          |
+------------------------+-------------+-------------+----------+
|                   5.00 |        7    |        7    |    14    |
|    7-9 HOURS           |       1.9%  |       1.6%  |    1.8%  |
|                        |        .7   |       -.7   |          |
+------------------------+-------------+-------------+----------+
|                   6.00 |        5    |        1    |     6    |
|    10-12 HOURS         |       1.4%  |        .2%  |     .8%  |
|                        |       2.3   |      -2.3   |          |
+------------------------+-------------+-------------+----------+
|                   7.00 |        1    |        0    |     1    |
|    13-15 HOURS         |        .3%  |        .0%  |     .1%  |
|                        |        .6   |       -.6   |          |
+------------------------+-------------+-------------+----------+
|                   8.00 |        4    |        1    |     5    |
|    OVER 15 HOURS       |       1.1%  |        .2%  |     .6%  |
|                        |       1.8   |      -1.8   |          |
+------------------------+-------------+-------------+----------+
|                 Column |      360    |      440    |   800    |
|                  Total |      45.0%  |      55.0%  |  100.0%  |
+------------------------+-------------+-------------+----------+

  Chi-Square     D.F.     Significance       Min E.F.      Cells with E.F.< 5
  ----------     ----     ------------       --------      ------------------
    24.05221       7          .0011              .450      6 OF    16 ( 37.5%)

Number of Missing Observations =      196
```

```
Crosstabulation:     Q24E2     TIME SPENT HIST/SOC STUD HW OUT OF SCH
                  By RTYPE     response type

                  Count     open           closed
  RTYPE->         Col Pct                             Row
                  Residual          .00        1.00   Total

  Q24E2                             73          101     174
                      1.00       20.1%        22.4%   21.4%
    NONE                          -4.7          4.7

                      2.00         161          150     311
    1 HR OR LESS                 44.4%        33.3%   38.3%
                                  22.1        -22.1

                      3.00          80          126     206
    2-3 HOURS                    22.0%        28.0%   25.3%
                                  12.0         12.0

                      4.00          32           45      77
    4-6 HOURS                     8.8%        10.0%    9.5%
                                  -2.4          2.4

                      5.00           8           15      23
    7-9 HOURS                     2.2%         3.3%    2.8%
                                  -2.3          2.3

                      6.00           5            6      11
    10-12 HOURS                   1.4%         1.3%    1.4%
                                    .1          -.1

                      7.00           3            1       4
    13-15 HOURS                    .8%          .2%     .5%
                                   1.2         -1.2

                      8.00           1            6       7
    OVER 15 HOURS                  .3%         1.3%     .9%
                                  -2.1          2.1

                    Column         363          450     813
                     Total       44.6%        55.4%  100.0%

 Chi-Square    D.F.     Significance     Min E.F.    Cells with E.F.< 5
 ----------    ----     ------------     --------    ------------------
   15.01624       7           .0358         1.786    5 OF    16 ( 31.3%)

Number of Missing Observations =      183
```

Crosstabulation:     Q24F1     TIME SPENT HW ALL OTHER SUBJ IN SCHOOL
By RTYPE     response type

| RTYPE-> Count<br>Col Pct<br>Residual | open<br>.00 | closed<br>1.00 | Row<br>Total |
|---|---|---|---|
| Q24A2<br>1.00<br>NONE | 175<br>38.0%<br>24.0 | 131<br>27.7%<br>-24.0 | 306<br>32.8% |
| 2.00<br>1 HR OR LESS | 153<br>33.2%<br>-39.0 | 236<br>49.9%<br>39.0 | 389<br>41.6% |
| 3.00<br>2-3 HOURS | 74<br>16.1%<br>10.8 | 54<br>11.4%<br>-10.8 | 128<br>13.7% |
| 4.00<br>4-6 HOURS | 26<br>5.6%<br>-1.1 | 29<br>6.1%<br>1.1 | 55<br>5.9% |
| 5.00<br>7-9 HOURS | 9<br>2.0%<br>.1 | 9<br>1.9%<br>-.1 | 18<br>1.9% |
| 6.00<br>10-12 HOURS | 10<br>2.2%<br>1.6 | 7<br>1.5%<br>-1.6 | 17<br>1.8% |
| 7.00<br>13-15 HOURS | 6<br>1.3%<br>1.1 | 4<br>.8%<br>-1.1 | 10<br>1.1% |
| 8.00<br>OVER 15 HOURS | 8<br>1.7%<br>2.6 | 3<br>.6%<br>-2.6 | 11<br>1.2% |
| Column<br>Total | 461<br>49.4% | 473<br>50.6% | 934<br>100.0% |

| Chi-Square | D.F. | Significance | Min E.F. | Cells with E.F.< 5 |
|---|---|---|---|---|
| 30.37793 | 7 | .0001 | 4.936 | 1 OF 16 ( 6.3%) |

Number of Missing Observations = 62

```
Crosstabulation:      Q24F2     TIME SPENT HW ALL OTHER SUBJ OUT OF SCHL
                   By RTYPE     response type

               Count      open           closed
  RTYPE->      Col Pct                                  Row
               Residual          .00            1.00    Total

   Q24F2                       128              98        226
                   1.00        27.5%           20.1%     23.7%
     NONE                      17.5           -17.5

                   2.00        163             190        353
     1 HR OR LESS              35.0%           39.0%     37.0%
                               -9.6             9.6

                   3.00        101             120        221
     2-3 HOURS                 21.7%           24.6%     23.2%
                               -7.1             7.1

                   4.00         51              40         91
     4-6 HOURS                 10.9%            8.2%      9.5%
                                6.5            -6.5

                   5.00         11              19         30
     7-9 HOURS                  2.4%            3.9%      3.1%
                               -3.7             3.7

                   6.00          3               8         11
     10-12 HOURS                 .6%            1.6%      1.2%
                               -2.4             2.4

                   7.00          2               5          7
     13-15 HOURS                 .4%            1.0%       .7%
                               -1.4             1.4

                   8.00          7               7         14
     OVER 15 HOURS              1.5%            1.4%      1.5%
                                 .2             -.2

                 Column        466             487        953
                  Total        48.9%           51.1%    100.0%

  Chi-Square    D.F.      Significance        Min E.F.     Cells with E.F.< 5
  ----------    ----      ------------        --------     ------------------
    14.24655       7          .0470             3.423      2 OF    16 ( 12.5%)

Number of Missing Observations =       43
```

```
Crosstabulation: Q36 Time spent on school-sponsored extracurricular activities
                  By RTYPE     response type

         Count    |open        |closed      |
RTYPE->  Residual |            |            | Row
         Col Pct  |         .00|        1.00| Total
------------------+------------+------------+--------
  Q36VT           |      121   |      155   |    276
            1.00  |     -9.1   |      9.1   |  29.8%
     NONE         |     27.8%  |     31.7%  |
------------------+------------+------------+--------
            2.00  |        0   |       67   |     67
     lt lhr per wk|    -31.6   |     31.6   |   7.2%
                  |       .0%  |     13.7%  |
------------------+------------+------------+--------
            3.00  |       90   |       99   |    189
     1-4 hrs wk   |       .9   |      -.9   |  20.4%
                  |     20.6%  |     20.2%  |
------------------+------------+------------+--------
            4.00  |       86   |       63   |    149
     5-9 hrs wk   |     15.8   |    -15.8   |  16.1%
                  |     19.7%  |     12.9%  |
------------------+------------+------------+--------
            5.00  |       91   |       88   |    179
     10-19 hrs wk |      6.6   |     -6.6   |  19.4%
                  |     20.9%  |     18.0%  |
------------------+------------+------------+--------
            6.00  |       48   |       17   |     65
     20+ hrs wk   |     17.4   |    -17.4   |   7.0%
                  |     11.0%  |      3.5%  |
------------------+------------+------------+--------
          Column  |      436   |      489   |    925
           Total  |     47.1%  |     52.9%  | 100.0%

 Chi-Square     D.F.     Significance       Min E.F.     Cells with E.F.< 5
 ----------     ----     ------------       --------     ------------------
   87.25189       5          .0000            30.638        None

Number of Missing Observations =       71
```

```
Crosstabulation:     Q39       # HRS/DAY WATCH TV/VIDEO TAPES
                  By RTYPE     response type

+----------------------+-------------+-------------+----------+
|             Count    | open        | closed      |          |
|  RTYPE->    Col Pct  |             |             | Row      |
|             Residual |         .00 |        1.00 | Total    |
+----------------------+-------------+-------------+----------+
|  Q39                 |          19 |          22 |       41 |
|                 1.00 |        4.1% |        4.5% |     4.3% |
|    DON'T WATCH TV    |        -1.0 |         1.0 |          |
+----------------------+-------------+-------------+----------+
|                 2.00 |           0 |          61 |       61 |
|    LT 1 HR           |         .0% |       12.4% |     6.3% |
|                      |       -29.7 |        29.7 |          |
+----------------------+-------------+-------------+----------+
|                 3.00 |         102 |         105 |      207 |
|    1-2 HOURS         |       21.8% |       21.3% |    21.5% |
|                      |         1.2 |        -1.2 |          |
+----------------------+-------------+-------------+----------+
|                 4.00 |          98 |         111 |      209 |
|    2-3 HOURS         |       20.9% |       22.5% |    21.7% |
|                      |        -3.8 |         3.8 |          |
+----------------------+-------------+-------------+----------+
|                 5.00 |         114 |         110 |      224 |
|    3-5 HOURS         |       24.4% |       22.3% |    23.3% |
|                      |         4.9 |        -4.9 |          |
+----------------------+-------------+-------------+----------+
|                 6.00 |          64 |          42 |      106 |
|     5-7 HOURS        |       13.7% |        8.5% |    11.0% |
|                      |        12.4 |       -12.4 |          |
+----------------------+-------------+-------------+----------+
|                 7.00 |          71 |          42 |      113 |
|    OVER 7 HRS        |       15.2% |        8.5% |    11.8% |
|                      |        16.0 |       -16.0 |          |
+----------------------+-------------+-------------+----------+
|               Column |         468 |         493 |      961 |
|                Total |       48.7% |       51.3% |   100.0% |
+----------------------+-------------+-------------+----------+

 Chi-Square    D.F.     Significance      Min E.F.     Cells with E.F.< 5
 ----------    ----     ------------      --------     ------------------
   73.55096      6          .0000           19.967        None

Number of Missing Observations =      35
```

```
Crosstabulation:      Q97      # CIGARETTES YOU USUALLY SMOKE PER DAY
                   By RTYPE     response type

            Count      open             closed
  RTYPE->   Col Pct                                       Row
            Residual           .00              1.00      Total

  Q97                          305               301        606
                   1.00      85.4%             82.0%      83.7%
                               6.2              -6.2
      I DON'T SMOKE

                   2.00          0                13         13
                               .0%              3.5%       1.8%
      LT 1 CIG/DAY            -6.4               6.4

                   3.00         29                26         55
                              8.1%              7.1%       7.6%
      1-5 CIGARETTES/D         1.9              -1.9

                   4.00         12                16         28
                              3.4%              4.4%       3.9%
      1/2 PACK/DAY            -1.8               1.8

                   5.00         11                10         21
                              3.1%              2.7%       2.9%
      1/2-2 PACKS/DAY           .6               -.6

                   6.00          0                 1          1
                               .0%               .3%        .1%
      2 PACKS/DAY OR M         -.5                .5

                 Column        357               367        724
                  Total      49.3%             50.7%     100.0%

 Chi-Square      D.F.       Significance        Min E.F.      Cells with E.F.< 5
 ----------      ----       ------------        --------      ------------------
   14.67377        5             .0119              .493      2 OF    12 ( 16.7%)

Number of Missing Observations =      272
```

```
Crosstabulation:     Q98       # OCCASIONS HAD ALCOHOL TO DRINK
                  By RTYPE     response type

            Count      open       closed
RTYPE->     Col Pct                          Row
            Residual        .00        1.00  Total
Q98                        77          49      126
                   1.00   25.5%       14.1%   19.4%
                          18.4       -18.4
   0 OCCASIONS

                   2.00    53          68      121
                          17.5%       19.6%   18.6%
   1-2 OCCASIONS          -3.3         3.3

                   3.00   118         106      224
                          39.1%       30.5%   34.5%
   3-19 OCCASIONS         13.8       -13.8

                   4.00    54         124      178
                          17.9%       35.7%   27.4%
   20+ OCCASIONS         -28.8         28.8

                 Column   302         347      649
                  Total   46.5%       53.5%  100.0%

Chi-Square    D.F.     Significance      Min E.F.    Cells with E.F.< 5
----------    ----     ------------      --------    ------------------
  33.29254      3          .0000          56.305       None

Number of Missing Observations =     347
```

Table 3.4.5 Analysis of number of responses to Q12 by response format and response order.

| TOTAL POPULATION |
|---|
| 1.50<br>( 996) |

| FORMAT | |
|---|---|
| Forced Choice | Mark All That Apply |
| 1.35<br>( 499) | 1.64<br>( 497) |

| ORDER | |
|---|---|
| Natural | Reverse |
| 1.43<br>( 498) | 1.57<br>( 498) |

| | ORDER | |
|---|---|---|
| FORMAT | Natural | Reverse |
| Forced Choice | 1.32<br>( 250) | 1.39<br>( 249) |
| Mark All That Apply | 1.54<br>( 248) | 1.75<br>( 249) |

| Source of Variation | Sum of Squares | DF | Mean Square | F | Signif. of F |
|---|---|---|---|---|---|
| Main Effects | 25.838 | 2 | 12.919 | 13.406 | .000 |
| FORMAT | 20.776 | 1 | 20.776 | 21.559 | .000 |
| ORDER | 5.020 | 1 | 5.020 | 5.209 | .025 |
| 2-way Interactions | 1.171 | 1 | 1.171 | 1.215 | .271 |
| FORMAT ORDER | 1.171 | 1 | 1.171 | 1.215 | .271 |
| Explained | 27.008 | 3 | 9.003 | 9.342 | .000 |
| Residual | 955.982 | 992 | .964 | | |
| Total | 982.991 | 995 | .988 | | |

**Table 3.4.6** **Analysis of number of responses to Q31 by response format and response order.**

| TOTAL POPULATION |
|---|
| 1.73<br>( 996) |

| FORMAT | |
|---|---|
| Forced Choice | Mark All That Apply |
| 1.52<br>( 499) | 1.95<br>( 497) |

| ORDER | |
|---|---|
| Natural | Reverse |
| 1.66<br>( 498) | 1.80<br>( 498) |

| | ORDER | |
|---|---|---|
| FORMAT | Natural | Reverse |
| Forced Choice | 1.46<br>( 250) | 1.58<br>( 249) |
| Mark All That Apply | 1.87<br>( 248) | 2.02<br>( 249) |

| Source of Variation | Sum of Squares | DF | Mean Square | F | Signif. of F |
|---|---|---|---|---|---|
| Main Effects | 49.197 | 2 | 24.598 | 14.030 | .000 |
| FORMAT | 44.416 | 1 | 44.416 | 25.334 | .000 |
| ORDER | 4.722 | 1 | 4.722 | 2.693 | .101 |
| 2-way Interactions | .094 | 1 | .094 | .054 | .817 |
| FORMAT ORDER | .094 | 1 | .094 | .054 | .817 |
| Explained | 49.290 | 3 | 16.430 | 9.371 | .000 |
| Residual | 1739.203 | 992 | 1.753 | | |
| Total | 1788.493 | 995 | 1.797 | | |

**Table 3.4.7. Analysis of number of responses to Q33 by response format and response order.**

| TOTAL POPULATION |
|---|
| 2.94<br>( 996) |

| FORMAT | |
|---|---|
| Forced Choice | Mark All That Apply |
| 2.79<br>( 499) | 3.08<br>( 497) |

| ORDER | |
|---|---|
| Natural | Reverse |
| 2.97<br>( 498) | 2.90<br>( 498) |

| | ORDER | |
|---|---|---|
| | Natural | Reverse |
| FORMAT<br>Forced Choice | 2.86<br>( 250) | 2.72<br>( 249) |
| Mark All That Apply | 3.09<br>( 248) | 3.08<br>( 249) |

| Source of Variation | Sum of Squares | DF | Mean Square | F | Signif. of F |
|---|---|---|---|---|---|
| Main Effects | 22.688 | 2 | 11.344 | 6.936 | .001 |
| FORMAT | 21.386 | 1 | 21.386 | 13.076 | .000 |
| ORDER | 1.322 | 1 | 1.322 | .809 | .369 |
| 2-way Interactions | 1.031 | 1 | 1.031 | .631 | .427 |
| FORMAT ORDER | 1.031 | 1 | 1.031 | .631 | .427 |
| Explained | 23.719 | 3 | 7.906 | 4.834 | .002 |
| Residual | 1622.421 | 992 | 1.636 | | |
| Total | 1646.141 | 995 | 1.654 | | |

**Table 3.4.8.** **Analysis of number of responses to Q49 by response format and response order.**

| TOTAL POPULATION |
|---|
| .26<br>( 996) |

| FORMAT | |
|---|---|
| Forced Choice | Mark All That Apply |
| .17<br>( 499) | .35<br>( 497) |

| ORDER | |
|---|---|
| Natural | Reverse |
| .24<br>( 498) | .29<br>( 498) |

| | ORDER | |
|---|---|---|
| | Natural | Reverse |
| FORMAT Forced Choice | .17<br>( 250) | .18<br>( 249) |
| Mark All That Apply | .31<br>( 248) | .39<br>( 249) |

| Source of Variation | Sum of Squares | DF | Mean Square | F | Signif. of F |
|---|---|---|---|---|---|
| Main Effects | 8.214 | 2 | 4.107 | 4.704 | .009 |
| FORM | 7.683 | 1 | 7.683 | 8.800 | .003 |
| ORDER | .523 | 1 | .523 | .599 | .439 |
| 2-way Interactions | .274 | 1 | .274 | .314 | .575 |
| FORMAT ORDER | .274 | 1 | .274 | .314 | .575 |
| Explained | 8.488 | 3 | 2.829 | 3.241 | .021 |
| Residual | 866.117 | 992 | .873 | | |
| Total | 874.605 | 995 | .879 | | |

**Table 3.4.9.** **Analysis of number of responses to Q113 by response format and response order.**

| TOTAL POPULATION |
|---|
| 1.12<br>( 996) |

| FORMAT | |
|---|---|
| Forced Choice | Mark All That Apply |
| 1.02<br>( 499) | 1.22<br>( 497) |

| ORDER | |
|---|---|
| Natural | Reverse |
| 1.07<br>( 498) | 1.18<br>( 498) |

| | ORDER | |
|---|---|---|
| | Natural | Reverse |
| FORMAT<br>Forced Choice | .96<br>( 250) | 1.08<br>( 249) |
| Mark All That Apply | 1.17<br>( 248) | 1.27<br>( 249) |

| Source of Variation | Sum of Squares | DF | Mean Square | F | Signif. of F |
|---|---|---|---|---|---|
| Main Effects | 12.905 | 2 | 6.452 | 2.642 | .072 |
| FORM | 9.868 | 1 | 9.868 | 4.040 | .045 |
| ORDER | 3.015 | 1 | 3.015 | 1.235 | .267 |
| 2-way Interactions | .026 | 1 | .026 | .011 | .917 |
| FORMAT ORDER | .026 | 1 | .026 | .011 | .917 |
| Explained | 12.931 | 3 | 4.310 | 1.765 | .152 |
| Residual | 2422.879 | 992 | 2.442 | | |
| Total | 2435.810 | 995 | 2.448 | | |

## APPENDIX C

### Documentation in Support of Analyses of Cognitive Test Data

Samples of Free Response Items

## Figure 3.10-2
## Form 1: Mathematics Free Response Items

### Question 1: Square Grid Problem

Purpose: Test understanding of area and perimeter

What is area of square shown? Sketch square of half the perimeter. Sketch square of half the area. Explain how you determined the length of the side.

| | |
|---|---|
| 1-nothing correct | 15% |
| 2-correct area only | 44% |
| 3-one correct sketch | 25% |
| 4-two correct sketches | 3% |
| 5-two correct sketches, plus correct explanation | 14% |

### Question 2: Balance Beam

Purpose: Simple word problem, algebra skills

Draw weights that will make the beam balance: equal weights; unequal weights; an additional weight; write a general formula/ explanation

| | |
|---|---|
| 1-nothing correct | 7% |
| 2-only equal weights correct | 36% |
| 3-unequal weights correct | 20% |
| 4-additional weight correct | 13% |
| 5-as above, explanation partially correct | 11% |
| 6-as above, explanation correct | 7% |
| x-inconsistent response | 5% |

### Question 3: Number System

Purpose: Test understanding of representation of numbers

A different way of representing numbers is defined. Draw the number of x's represented. Convert to and from the pattern described. Extrapolate to the next level and explain.

| | |
|---|---|
| 1-nothing correct | 20% |
| 2-only group of x's correct | 4% |
| 3-conversion to or from correct | 23% |
| 4-conversion to and from correct | 48% |
| 5-extrapolation and explanation essentially correct | 5% |

### Question 4: Equilateral and Isosceles Triangles

Purpose: Test understanding of area and perimeter

A. Explain why a generalization about perimeter and area of equilateral triangles is true.
B. Does the same generalization apply to isosceles triangles? Explain.

| | |
|---|---|
| 1-nothing correct | 62% |
| 2-some understanding, Part A | 18% |
| 3-A correct, B incorrect | 8% |
| 4-A correct, some understanding of Part B | 6% |
| 5-A and B correct | 3% |
| x-inconsistent response | 4% |

## Table 3.10.3
## Form 2: Mathematics Free Response Items

### Question 1: Train Schedule

Purpose: Application to a real life situation

Read an entry in the train schedule. Which train is fastest? What time must one leave home for an appointment, given travel times to and from both train stations? Write a general formula adjusting the schedule for later departures and 10% longer travel time.

| | |
|---|---|
| 1=nothing correct | 4% |
| 2=one of parts A-C correct | 27% |
| 3=two of parts A-C correct | 29% |
| 4=all three A-C correct | 30% |
| 5=at least two A-C correct and formula partly correct | 7% |
| 6=at least two A-C correct and formula correct | 4% |

### Question 2: Area of Figure Made of Rectangles

Purpose: Decomposition of figure and computation of area.

Compute area of simple figure with decomposition lines marked. Decompose another figure into rectangles and compute area. Compute area of a complex figure.

| | |
|---|---|
| 1=nothing correct | 20% |
| 2=simple figure correct | 20% |
| 3=second level figure correct, complex figure incorrect | 29% |
| 4=second level correct, complex figure partially correct | 10% |
| 5=second level figure and complex figure correct | 18% |
| x=inconsistent response | 3% |

### Question 3: Dice Problem

Purpose: Probability

Give an example of a winning combination. How many combinations are possible? What is the probability of winning?

| | |
|---|---|
| 1=nothing correct | 14% |
| 2=example correct only | 40% |
| 3=# combinations correct | 37% |
| 4=partly correct probability | 1% |
| 5=correct probability | 2% |
| 6=correct probability and explanation | 7% |

### Question 4: Car Stopping Distance

Purpose: Graph reading, real life application

Interpret graphs of Driver Reaction Distance and Car Braking Distance. Estimate total stopping distance. Find shortest possible distance to collision; explain why it could have been longer.

| | |
|---|---|
| 1=nothing correct | 15% |
| 2=graph reading correct only | 48% |
| 3=stopping distance partly correct | 1% |
| 4=stopping distance correct | 8% |
| 5=distance to collision partly correct | 16% |
| 6=distance to collision correct | 12% |

## Table 3.10.4
## Form 3: Science Free Response Items

### Question 1A: Pulling Rock (Figural Response)

Purpose: Short answer

Two children are pulling on a rock at a 90 degree angle from each other. Draw arrow showing direction the rock will move.

| | |
|---|---|
| 1=out of quadrant | 5% |
| 2=right quadrant, not 45 degrees | 4% |
| 3=right quadrant, 45% angle | 91% |

### Question 1B: Blood Flow (Figural Response)

Purpose: Short answer

Diagram of part of circulatory system has lines in superior vena cava, aorta, left atrium, right ventricle, and blood vessel from gut. Draw arrow head on each line to show direction of blood flow.

| | |
|---|---|
| 1=at least one arrowhead drawn, nothing correct | 10% |
| 2=one correct | 13% |
| 3=two correct | 22% |
| 4=three correct | 17% |
| 5=four correct | 20% |
| 6=all five correct | 20% |

### Question 1C: Wind Direction (Figural Response)

Purpose: Short answer

Map shows a high pressure and a low pressure area. Draw arrow showing wind direction.

| | |
|---|---|
| 1=incorrect arrow drawn | 91% |
| 2=correct direction | 9% |

### Question 1D: Rotating Ball on a String (Figural Response)

Purpose: Short answer

Diagram shows ball being swung in a circular path on the end of a string. Draw arrow to show path ball will take when released.

| | |
|---|---|
| 1=incorrect arrow drawn | 66% |
| 2=correct direction | 34% |

### Question 1E: Radioactive Decay (Figural Response)

Purpose: Short answer

Radioactive substance has a half life of 2 years. On the graph plot decay curve over 8 years.

| | |
|---|---|
| 1=incorrect | 59% |
| 2=partially correct | 12% |
| 3=correct | 28% |

## Table 3.10.4
## Form 3: Science (Continued)

### Question 2: Eclipses

Purpose: Test understanding of positions of earth, moon, and sun responsible for eclipses.

Draw diagram of relative positions of earth, moon, and sun during a solar eclipse. Draw diagram of lunar eclipse. Explain why lunar eclipse can be seen in many more locations on earth.

| | |
|---|---|
| 1=nothing correct | 31% |
| 2=part A or B correct | 39% |
| 3=part A and B correct | 20% |
| 4=A and B correct, partial explanation | 1% |
| 5=A and B correct, correct explanation | 9% |

### Question 3: Colliding Railroad Cars

Purpose: Test knowledge and application of law of conservation of momentum.

A railroad car traveling 12 m/s hits another car of twice its mass at rest. Calculate common speed after impact and write equations used.

| | |
|---|---|
| 1=nothing correct | 79% |
| 2=correct answer, equation absent, irrelevant, or incorrect | 13% |
| 3=correct answer, correct equation | 8% |

### Question 4: Heating Curve

Purpose: Test understanding of absorption of heat.

Beaker containing water and ice is heated. Temperature of mixture over time is shown on a graph. Explain why first section of graph is horizontal; why next section slopes upward; why last section is horizontal.

| | |
|---|---|
| 1=nothing correct | 14% |
| 2=reference to temperature only | 37% |
| 3=mention of melting or boiling for horizontal sections | 36% |
| 4=includes melting/boiling, temperature, and absorption of heat; or mention of kinetic/potential energy change | 13% |

## Table 3.10.5
## Form 4: Science

### Question 1: Position of Earth and Sun

Purpose: Test understanding of relationship between seasons and earth's tilt and position

Draw a diagram of positions of Sun and Earth on June 21 and December 21. Explain how and why seasons differ for these two positions.

| | |
|---|---|
| 1=nothing correct | 50% |
| 2=any one of (axis, positions, distance) correct | 17% |
| 3=as above, plus partly correct explanation | 11% |
| 4=two or three features correct | 18% |
| 5=as above, plus complete and correct explanation | 4% |

### Question 2: Nuclear vs. Fossil Fuels

Purpose: Test awareness of issues regarding alternative fuel sources.

List and explain at least 3 advantages/disadvantages of nuclear vs. fossil fuels.

| | |
|---|---|
| 1=no valid statements, too many incorrect statements, or only emotional rather than rational reasons given | 20% |
| 2=one valid response | 13% |
| 3=two valid responses | 16% |
| 4=three or more valid responses | 30% |
| 5=one point was added if there were any elaborations (one subtracted if any incorrect statements were made) | 21% |

### Question 3: Rabbit/Wolf Populations

Purpose: Test understanding of predator-prey relationship.

Graph has rabbit population over time. Draw a curve showing wolf population and explain the shape and location.

| | |
|---|---|
| 1=nothing correct | 23% |
| 2=any one of (phase,lag,height) correct in drawing | 19% |
| 3=two correct features | 30% |
| 4=three correct features | 27% |
| 5=one point was added to any score for a correct explanation (one subtracted if any incorrect statements were made) | 1% |

### Question 4: Stoichiometry

Purpose: Test whether student can balance a chemical equation.

Write a balanced equation for the reaction described, and show all calculations.

| | |
|---|---|
| 1=nothing correct | 27% |
| 2=unbalanced equation | 28% |
| 3=balanced equation, unreduced coefficients | 9% |
| 4=balanced equation, smallest possible coefficients | 17% |
| 5=any equation, plus one or more correct calculations | 15% |
| 6=balanced equation, smallest coefficients, three or more correct calculations | 5% |

## Table 3.10.6
## Form 5: Reading Comprehension

### Question 1: Forget that House?

Purpose: Convert description into drawing

The reading passage describes a house, with details about the appearance of features such as its height, roof, towers, number and placement of windows and doors. Draw the house and label the parts.

| | |
|---|---|
| 1=nothing correct | 1% |
| 2=basic outline incorrect | 49% |
| 3=basic outline correct but nothing else | 2% |
| 4=basic outline plus at least two of (roof,windows,doors) | 3% |
| 5=as above, plus labelling or tower windows | 23% |
| 6=all features essentially correct | 23% |

### Question 2: Producing Rice

Purpose: Test ability to select and order relevant important points.

The passages describes the steps involved in producing rice for the table. List and number the steps taken after harvesting to produce white rice.

| | |
|---|---|
| 1=nothing correct | 3% |
| 2=only one main point, or something out of order | 16% |
| 3=at least two points, correct order | 22% |
| 4=at least three points in order, $\leq 2$ additional steps | 20% |
| 5=four main points, as above | 30% |
| 6=as above, but no additional or irrelevant steps | 8% |
| 7=as above, optional step recognized | 2% |

### Question 3: Letter of Recommendation

Purpose: Testing the ability to draw inferences, to understand implied meaning.

Letter of recommendation ostensibly praises candidate but hints at problems. Would you hire him? Give reasons for decision.

| | |
|---|---|
| 1=nothing correct | 1% |
| 2=inadequate explanation and unclear writing; or no valid reasons; or $\geq 2$ invalid reasons | 24% |
| 3=$\geq 1$ valid reason, $\leq 1$ invalid | 1% |
| 4=as above, plus clear writing | 19% |
| 5=as above, plus hidden meaning | 12% |
| 6=score 4, plus logical expl. and $\geq 2$ acceptable reasons | 13% |
| 7=score 6, plus hidden meaning | 28% |
| 8=$\geq 3$ acceptable reasons, no unacceptable, hidden meaning, clear and logical explanation | 1% |

### Question 4: Health-Glo for Your Heart

Purpose: Manipulative language.

An advertisement for a health food product claims that its main ingredient is eaten by Pacific pearl divers, who have a low incidence of heart disease. What is the advertiser asking you to believe? What other explanation can you give for the divers' good health?

| | |
|---|---|
| 1=nothing correct | 4% |
| 2=some insight but inadequate response | 18% |
| 3=valid but limited response | 65% |
| 4=awareness of hidden assumption | 11% |
| 5=hidden assumption plus combination of accurate factors in explanation | 2% |

### Table 3.10.8
### Field Test Item Statistics
### Reading Comprehension Test

N=411

| Item # | 1991 R-Biserial | Proportion Correct (P+) 1991 | Proportion Correct (P+) 1989 | 1991-1989 | Adj. P+ 1991 | Source | Content |
|---|---|---|---|---|---|---|---|
| 1-1 | .56 | .49 | .46 | .03 | .41 | NELS | Biography |
| 1-2 | .66 | .43 | .36 | .07 | .35 | NELS | Biography |
| 1-3 | .49 | .30 | .25 | .05 | .22 | NELS | Biography |
| 1-4 | .60 | .45 | .38 | .07 | .38 | NELS | Biography |
| 1-5 | .57 | .34 | .28 | .06 | .26 | NELS | Biography |
| 1-6 | .36 | .40 | .34 | .06 | .33 | NELS | Biography |
| 1-7 | .43 | .83 | .74 | .09 | .75 | 3IBR | Letter |
| 1-8 | .59 | .87 | .76 | .11 | .79 | 3IBR | Letter |
| 1-9 | .59 | .75 | .68 | .07 | .67 | 3IBR | Letter |
| 1-10 | .59 | .73 | .60 | .13 | .65 | 3IBR | Letter |
| 1-11 | .53 | .73 | .62 | .11 | .66 | 3IBR | Letter |
| 1-12 | .41 | .60 | .57 | .03 | .52 | 3IBR | Letter |
| 1-13 | .62 | .86 | .78 | .08 | .79 | 3IBR | Letter |
| 1-14 | .58 | .78 | .67 | .11 | .70 | 3IBR | Letter |
| 1-15 | .45 | .55 | | | .48 | New Item | Comparisons |
| 1-16 | .59 | .74 | | | .66 | New Item | Comparisons |
| 1-17 | .54 | .67 | | | .59 | New Item | Comparisons |
| 1-18 | .63 | .59 | | | .51 | New Item | Comparisons |
| 1-19 | .28 | .20 | | | .12 | New Item | Comparisons |
| 1-20 | .72 | .58 | | | .50 | New Item | Comparisons |
| 1-21 | .47 | .60 | | | .52 | New Item | Comparisons |
| 1-22 | .63 | .63 | | | .55 | New Item | Comparisons |
| 2-1 | .46 | .33 | | | .25 | New Item | Comparisons |
| 2-2 | .65 | .66 | | | .58 | New Item | Comparisons |
| 2-3 | .46 | .43 | | | .36 | New Item | Comparisons |
| 2-4 | .78 | .84 | | | .77 | New Item | Comparisons |
| 2-5 | .71 | .71 | | | .64 | New Item | Comparisons |
| 2-6 | .38 | .39 | | | .32 | New Item | Comparisons |
| 2-7 X | .20 | .28 | | | .20 | New Item | Comparisons |
| 2-8 | .39 | .32 | | | .24 | New Item | Comparisons |
| 2-9 | .56 | .44 | | | .36 | New Item | Comparisons |
| 2-10 X | .18 | .41 | | | .33 | New Item | Comparisons |
| 2-11 | .51 | .66 | | | .58 | New Item | Comparisons |
| 2-12 | .54 | .55 | | | .47 | New Item | Comparisons |
| 2-13 | .47 | .43 | | | .35 | New Item | Comparisons |
| 2-14 | .57 | .66 | | | .58 | New Item | Comparisons |
| 2-15 | .42 | .35 | | | .28 | New Item | Comparisons |
| 2-16 | .28 | .43 | | | .36 | New Item | Comparisons |
| 2-17 | .41 | .39 | | | .31 | New Item | Comparisons |
| 2-18 | .65 | .70 | | | .62 | New Item | Comparisons |
| Col Mean | .51 | .55 | .53 | .08 | .48 | | |
| Col S.D. | .13 | .18 | .18 | .03 | .18 | | |

## Table 3.10.9
## Field Test Item Statistics
## Mathematics Test

N=811

| Item # | | 1991 R-Biserial | Proportion Correct (P+) 1991 | Proportion Correct (P+) 1989 | 1991-1989 | Adj. P+ 1991 | Source | Content |
|---|---|---|---|---|---|---|---|---|
| 1-1 | | .59 | .31 | | | .29 | SIMS | Algebra |
| 1-2 | | .66 | .50 | | | .48 | NAEP | Geometry |
| 1-3 | | .50 | .61 | .57 | .04 | .59 | NAEP | Arithmetic |
| 1-4 | X | .01 | .17 | | | .15 | SIMS | Arithmetic |
| 1-5 | | .51 | .75 | | | [illegible] | NAEP | Geometry |
| 1-6 | | .44 | .71 | .65 | .06 | .6[illegible] | NAEP | Sets/Countng |
| 1-7 | | .58 | .53 | | | .50 | NAEP | Geometry |
| 1-8 | | .56 | .65 | .60 | .05 | .63 | NELS | Arithmetic |
| 1-9 | | .50 | .29 | .33 | -.04 | .27 | NAEP | AG/Func/PreCalc |
| 1-10 | | .44 | .46 | .43 | .03 | .44 | NAEP | Stat/Prob |
| 1-11 | | .60 | .41 | | | .39 | NAEP | Geometry |
| 1-12 | | .43 | .40 | .32 | .08 | .38 | NELS | Sets/Countng |
| 1-13 | | .50 | .71 | .73 | -.02 | .69 | NAEP | Stat/Prob |
| 1-14 | | .36 | .32 | | | .29 | SIMS | Geometry |
| 1-15 | | .39 | .50 | | | .47 | SIMS | Geometry |
| 1-16 | | .54 | .15 | | | .12 | SIMS | Stat/Prob |
| 1-17 | | .54 | .19 | .16 | .03 | .16 | NELS | Geometry |
| 1-18 | | .65 | .09 | .11 | -.02 | .07 | NELS | Geometry |
| 1-19 | X | .08 | .29 | | | .27 | SIMS | Geometry |
| 1-20 | | .36 | .20 | | | .18 | SIMS | Geometry |
| 2-1 | X | .24 | .28 | | | .26 | SIMS | Algebra |
| 2-2 | | .39 | .19 | | | .17 | SIMS | AG/Func/PreCalc |
| 2-3 | | .71 | .21 | | | .19 | SIMS | AG/Func/PreCalc |
| 2-4 | | .47 | .40 | | | .37 | SIMS | Sets/Countng |
| 2-5 | | .63 | .37 | | | .34 | SIMS | AG/Func/PreCalc |
| 2-6 | | .29 | .30 | | | .28 | SIMS | AG/Func/PreCalc |
| 2-7 | | .53 | .34 | | | .32 | NAEP | Geometry |
| 2-8 | | .57 | .26 | | | .23 | SIMS | AG/Func/PreCalc |
| 2-9 | X | .12 | .10 | | | .07 | SIMS | Sets/Countng |
| 2-10 | | .34 | .36 | | | .33 | SIMS | Sets/Countng |
| 2-11 | | .54 | .19 | | | .16 | NAEP | AG/Func/PreCalc |
| 2-12 | | .61 | .26 | | | .23 | SIMS | AG/Func/PreCalc |
| Col Mean | | .46 | .36 | .43 | .03 | .33 | | |
| Col S.D. | | .17 | .18 | .21 | .04 | .18 | | |

## Table 3.10.10
## Field Test Item Statistics
## Science Test

N=777

| Item # | | 1991 R-Biserial | Proportion Correct (P+) 1991 | Proportion Correct (P+) 1989 | 1991-1989 | Adj. P+ 1991 | Source | Content |
|---|---|---|---|---|---|---|---|---|
| 1 | | .64 | .76 | .82 | -.06 | .76 | hsb | Earth Science |
| 2 | | .70 | .85 | .87 | -.02 | .85 | hsb | Life Science |
| 3 | | .63 | .64 | .65 | -.01 | .64 | NAEP | Chemistry |
| 4 | | .57 | .62 | .64 | -.02 | .62 | NAEP | Life Science |
| 5 | | .49 | .56 | .54 | .02 | .56 | NAEP | Earth Science |
| 6 | | .40 | .49 | .41 | .08 | .49 | NAEP | Life Science |
| 7 | | .58 | .46 | .42 | .04 | .46 | hsb | Sci Method |
| 8 | | .59 | .53 | .55 | -.02 | .53 | NAEP | Physics |
| 9 | | .60 | .55 | .57 | -.02 | .55 | NAEP | Earth Science |
| 10 | | .42 | .55 | .55 | .00 | .55 | NAEP | Life Science |
| 11 | | .41 | .11 | | | .11 | New Item | Physics |
| 12 | | .33 | .27 | | | .27 | New Item | Physics |
| 13 | | .39 | .19 | | | .19 | New Item | Earth Science |
| 14 | | .43 | .35 | | | .35 | New Item | Physics |
| 15 | | .23 | .32 | | | .32 | New Item | Physics |
| 16 | | .40 | .54 | | | .54 | New Item | Physics |
| 17 | X | .10 | .21 | | | .21 | New Item | Chemistry |
| 18 | | .25 | .32 | | | .32 | New Item | Chemistry |
| 19 | | .26 | .14 | | | .14 | New Item | Physics |
| 20 | | .44 | .21 | | | .21 | New Item | Chemistry |
| 21 | | .31 | .16 | | | .16 | New Item | Chemistry |
| 22 | | .30 | .38 | | | .38 | New Item | Physics |
| 23 | | .52 | .39 | | | .39 | New Item | Chemistry |
| 24 | X | .09 | .30 | | | .30 | New Item | Chemistry |
| Col Mean | | .42 | .41 | .60 | .00 | .41 | | |
| Col S.D. | | .16 | .20 | .14 | .04 | .20 | | |

## Table 3.10.11
## Field Test Item Statistics
## History Test

N=769

| Item # | 1991 R-Biserial | Proportion Correct (P+) 1991 | Proportion Correct (P+) 1989 | 1991-1989 | Adj. P+ 1991 | Source | Content |
|---|---|---|---|---|---|---|---|
| 1 | .59 | .74 | | | .77 | NAEP | Geography |
| 2 | .62 | .86 | .90 | -.04 | .89 | NAEP | History |
| 3 | .48 | .75 | | | .78 | NAEP | Citizenship |
| 4 | .58 | .56 | | | .59 | NAEP | Geography |
| 5 | .62 | .29 | .30 | -.01 | .32 | NELS | History |
| 6 | .56 | .57 | | | .60 | NAEP | History |
| 7 | .46 | .47 | | | .50 | NAEP | Citizenship |
| 8 | .66 | .63 | .70 | -.07 | .66 | NAEP | Citizenship |
| 9 | .62 | .56 | .61 | -.05 | .59 | NAEP | Citizenship |
| 10 | .55 | .48 | | | .51 | NAEP | Geography |
| 11 | .55 | .39 | .42 | -.03 | .42 | NAEP | History |
| 12 | .31 | .46 | .43 | .03 | .48 | NAEP | History |
| 13 | .67 | .93 | .96 | -.03 | .96 | NAEP | Citizenship |
| 14 | .59 | .88 | .89 | -.01 | .91 | NAEP | Geography |
| 15 | .55 | .49 | | | .52 | NAEP | History |
| 16 X | .21 | .46 | | | .49 | NAEP | Geography |
| 17 | .47 | .31 | | | .34 | NAEP | History |
| 18 | .35 | .32 | | | .35 | NAEP | Citizenship |
| 19 | .52 | .33 | | | .36 | NAEP | Citizenship |
| 20 | .61 | .76 | .82 | -.06 | .79 | HSB | Citizenship |
| 21 | .40 | .32 | | | .34 | NAEP | History |
| 22 | .36 | .26 | .25 | .01 | .29 | NAEP | History |
| 23 | .41 | .49 | .54 | -.05 | .52 | NAEP | History |
| 24 | .40 | .31 | | | .33 | NAEP | Geography |
| 25 | .35 | .25 | | | .28 | NAEP | Citizenship |
| Col Mean | .50 | .51 | .62 | -.03 | .54 | | |
| Col S.D. | .12 | .20 | .24 | .03 | .20 | | |

## Table 3.10.12

### Form 1: Mathematics
### Percent of Test Items Omitted By Subgroup

| | Total | Male | Female | Black | White | Hispanic |
|---|---|---|---|---|---|---|
| Number of Cases | 405 | 215 | 184 | 58 | 222 | 68 |
| Multiple Choice Section 1 | 6% | 6% | 5% | 6% | 5% | 6% |
| Multiple Choice Section 2 | 6% | 7% | 6% | 8% | 6% | 6% |
| All Multiple Choice | 6% | 6% | 5% | 7% | 6% | 6% |
| Free Response Question 1 | 4% | 3% | 5% | 10% | 2% | 6% |
| % Hard/Too Hard | 56% | 50% | 67% | 67% | 40% | 75% |
| % Easy/Right | 13% | 0% | 11% | 0% | 20% | 0% |
| % No Difficulty Reaction | 31% | 50% | 22% | 33% | 40% | 25% |
| Free Response Question 2 | 7% | 7% | 7% | 14% | 5% | 12% |
| % Hard/Too Hard | 46% | 40% | 54% | 88% | 36% | 25% |
| % Easy/Right | 0% | 0% | 0% | 0% | 0% | 0% |
| % No Difficulty Reaction | 54% | 60% | 46% | 13% | 64% | 75% |
| Free Response Question 3 | 11% | 14% | 8% | 17% | 10% | 10% |
| % Hard/Too Hard | 46% | 40% | 60% | 70% | 43% | 29% |
| % Easy/Right | 11% | 17% | 0% | 0% | 9% | 14% |
| % No Difficulty Reaction | 43% | 43% | 40% | 30% | 48% | 57% |
| Free Response Question 4 | 34% | 33% | 35% | 57% | 27% | 44% |
| % Hard/Too Hard | 54% | 56% | 52% | 73% | 46% | 57% |
| % Easy/Right | 11% | 8% | 14% | 12% | 7% | 17% |
| % No Difficulty Reaction | 35% | 36% | 34% | 15% | 47% | 27% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.13

### Form 2: Mathematics
### Percent of Test Items Omitted By Subgroup

| | Total | Male | Female | Black | White | Hispanic |
|---|---|---|---|---|---|---|
| Number of Cases | 431 | 199 | 224 | 58 | 243 | 69 |
| Multiple Choice Section 1 | 6% | 5% | 6% | 7% | 6% | 5% |
| Multiple Choice Section 2 | 7% | 6% | 7% | 7% | 7% | 8% |
| All Multiple Choice | 6% | 5% | 6% | 7% | 6% | 6% |
| Free Response Question 1 | 2% | 2% | 1% | 3% | 1% | 1% |
| % Hard/Too Hard | 43% | 50% | 50% | 50% | 33% | 0% |
| % Easy/Right | 29% | 25% | 0% | 50% | 33% | 0% |
| % No Difficulty Reaction | 29% | 25% | 50% | 0% | 33% | 100% |
| Free Response Question 2 | 6% | 8% | 4% | 12% | 3% | 6% |
| % Hard/Too Hard | 58% | 53% | 75% | 57% | 63% | 75% |
| % Easy/Right | 8% | 0% | 13% | 0% | 13% | 0% |
| % No Difficulty Reaction | 33% | 47% | 13% | 43% | 25% | 25% |
| Free Response Question 3 | 7% | 9% | 4% | 9% | 7% | 6% |
| % Hard/Too Hard | 30% | 28% | 30% | 20% | 13% | 50% |
| % Easy/Right | 3% | 0% | 0% | 0% | 6% | 0% |
| % No Difficulty Reaction | 67% | 72% | 70% | 80% | 81% | 50% |
| Free Response Question 4 | 14% | 17% | 12% | 17% | 13% | 16% |
| % Hard/Too Hard | 17% | 9% | 27% | 30% | 13% | 9% |
| % Easy/Right | 3% | 6% | 0% | 10% | 0% | 0% |
| % No Difficulty Reaction | 80% | 85% | 73% | 60% | 88% | 91% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.14

### Form 3: Science and History/Citizenship/Geography
### Percent of Test Items Omitted By Subgroup

| | Total | Male | Female | Black | White | Hispanic |
|---|---|---|---|---|---|---|
| Number of Cases | 411 | 188 | 211 | 77 | 208 | 78 |
| Multiple Choice Science | 5% | 4% | 5% | 5% | 4% | 5% |
| Multiple Choice HCG | 3% | 3% | 3% | 5% | 2% | 2% |
| All Multiple Choice | 4% | 3% | 4% | 5% | 3% | 3% |
| Free Response Question 1 | 11% | 8% | 12% | 17% | 7% | 10% |
| % Hard/Too Hard | 72% | 64% | 76% | 64% | 71% | 77% |
| % Easy/Right | 22% | 26% | 21% | 31% | 25% | 14% |
| % No Difficulty Reaction | 6% | 10% | 2% | 5% | 4% | 9% |
| Free Response Question 2 | 11% | 6% | 16% | 21% | 7% | 12% |
| % Hard/Too Hard | 72% | 55% | 82% | 94% | 71% | 67% |
| % Easy/Right | 9% | 18% | 6% | 0% | 14% | 11% |
| % No Difficulty Reaction | 19% | 27% | 12% | 6% | 14% | 22% |
| Free Response Question 3 | 43% | 40% | 46% | 51% | 38% | 45% |
| % Hard/Too Hard | 71% | 68% | 77% | 85% | 73% | 74% |
| % Easy/Right | 10% | 12% | 8% | 5% | 14% | 6% |
| % No Difficulty Reaction | 19% | 20% | 14% | 10% | 13% | 20% |
| Free Response Question 4 | 24% | 22% | 25% | 35% | 20% | 21% |
| % Hard/Too Hard | 42% | 51% | 40% | 48% | 55% | 31% |
| % Easy/Right | 2% | 0% | 4% | 7% | 0% | 0% |
| % No Difficulty Reaction | 56% | 49% | 57% | 44% | 45% | 69% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

### Table 3.10.15

### Form 4: Science and History/Citizenship/Geography
### Percent of Test Items Omitted By Subgroup

| | Total | Male | Female | Black | White | Hispanic |
|---|---|---|---|---|---|---|
| Number of Cases | 367 | 176 | 187 | 53 | 195 | 73 |
| Multiple Choice Section 1 | 5% | 6% | 5% | 7% | 5% | 4% |
| Multiple Choice Section 2 | 4% | 4% | 3% | 5% | 3% | 3% |
| All Multiple Choice | 4% | 5% | 4% | 6% | 4% | 4% |
| Free Response Question 1 | 19% | 18% | 20% | 25% | 19% | 15% |
| % Hard/Too Hard | 73% | 66% | 79% | 62% | 71% | 82% |
| % Easy/Right | 11% | 13% | 11% | 15% | 13% | 9% |
| % No Difficulty Reaction | 16% | 22% | 11% | 23% | 16% | 9% |
| Free Response Question 2 | 37% | 34% | 40% | 60% | 33% | 29% |
| % Hard/Too Hard | 69% | 62% | 74% | 66% | 70% | 67% |
| % Easy/Right | 15% | 22% | 11% | 16% | 14% | 24% |
| % No Difficulty Reaction | 16% | 17% | 15% | 19% | 16% | 10% |
| Free Response Question 3 | 18% | 19% | 18% | 38% | 14% | 12% |
| % Hard/Too Hard | 60% | 55% | 67% | 55% | 64% | 78% |
| % Easy/Right | 3% | 3% | 3% | 10% | 0% | 0% |
| % No Difficulty Reaction | 37% | 42% | 30% | 35% | 36% | 22% |
| Free Response Question 4 | 59% | 65% | 53% | 85% | 54% | 59% |
| % Hard/Too Hard | 71% | 71% | 71% | 64% | 75% | 72% |
| % Easy/Right | 5% | 5% | 5% | 11% | 5% | 0% |
| % No Difficulty Reaction | 24% | 24% | 24% | 24% | 21% | 28% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.16

### Form 5: Reading Comprehension
### Percent of Test Items Omitted By Subgroup

| | Total | Male | Female | Black | White | Hispanic |
|---|---|---|---|---|---|---|
| Number of Cases | 415 | 201 | 210 | 59 | 229 | 84 |
| Multiple Choice Section 1 | 2% | 1% | 2% | 3% | 1% | 3% |
| Multiple Choice Section 2 | 3% | 2% | 3% | 7% | 2% | 2% |
| All Multiple Choice | 2% | 2% | 3% | 5% | 1% | 3% |
| Free Response Question 1 | 6% | 6% | 4% | 10% | 3% | 5% |
| % Hard/Too Hard | 48% | 46% | 50% | 50% | 57% | 25% |
| % Easy/Right | 39% | 46% | 25% | 50% | 29% | 50% |
| % No Difficulty Reaction | 13% | 8% | 25% | 0% | 14% | 25% |
| Free Response Question 2 | 5% | 6% | 3% | 12% | 2% | 5% |
| % Hard/Too Hard | 23% | 31% | 14% | 14% | 40% | 25% |
| % Easy/Right | 36% | 46% | 14% | 43% | 40% | 0% |
| % No Difficulty Reaction | 41% | 23% | 71% | 43% | 20% | 75% |
| Free Response Question 3 | 6% | 9% | 2% | 14% | 3% | 5% |
| % Hard/Too Hard | 17% | 21% | 0% | 13% | 29% | 0% |
| % Easy/Right | 26% | 32% | 0% | 38% | 29% | 0% |
| % No Difficulty Reaction | 57% | 47% | 100% | 50% | 43% | 100% |
| Free Response Question 4 | 12% | 12% | 10% | 12% | 9% | 15% |
| % Hard/Too Hard | 6% | 4% | 9% | 14% | 0% | 15% |
| % Easy/Right | 2% | 4% | 0% | 0% | 0% | 0% |
| % No Difficulty Reaction | 92% | 92% | 91% | 86% | 100% | 85% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.17

### Form 1: Mathematics
### Percent of Test Items Omitted By Coursework in Mathematics

| | Total | At least 1 year Algebra | At least 1 year Geometry | At least 1/2 year Trig | At least 1 year An. Geom or Calc |
|---|---|---|---|---|---|
| Number of Cases | 405 | 338 | 282 | 188 | 75 |
| Multiple Choice Section 1 | 6% | 5% | 5% | 5% | 5% |
| Multiple Choice Section 2 | 6% | 6% | 6% | 6% | 5% |
| All Multiple Choice | 6% | 6% | 5% | 6% | 5% |
| Free Response Question 1 | 4% | 2% | 1% | 1% | 0% |
| % Hard/Too Hard | 56% | 50% | 25% | 100% | 0% |
| % Easy/Right | 13% | 0% | 0% | 0% | 0% |
| % No Difficulty Reaction | 31% | 50% | 75% | 0% | 0% |
| Free Response Question 2 | 7% | 5% | 5% | 4% | 4% |
| % Hard/Too Hard | 46% | 44% | 36% | 29% | 67% |
| % Easy/Right | 0% | 0% | 0% | 0% | 0% |
| % No Difficulty Reaction | 54% | 56% | 64% | 71% | 33% |
| Free Response Question 3 | 11% | 10% | 9% | 9% | 11% |
| % Hard/Too Hard | 46% | 39% | 29% | 18% | 25% |
| % Easy/Right | 11% | 6% | 8% | 12% | 13% |
| % No Difficulty Reaction | 43% | 55% | 63% | 71% | 63% |
| Free Response Question 4 | 34% | 31% | 25% | 22% | 21% |
| % Hard/Too Hard | 54% | 53% | 41% | 27% | 13% |
| % Easy/Right | 11% | 13% | 16% | 12% | 19% |
| % No Difficulty Reaction | 35% | 35% | 43% | 61% | 69% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.18

### Fo[illegible]: Mathematics
### Percent of Test Items Omitted By Coursework in Mathematics

| | Total | At least 1 year Algebra | At least 1 year Geometry | At least 1/2 year Trig | At least 1 year An. Geom or Calc |
|---|---|---|---|---|---|
| Number of Cases | 431 | 354 | 312 | 192 | 84 |
| Multiple Choice Section 1 | 6% | 5% | 5% | 5% | 5% |
| Multiple Choice Section 2 | 7% | 6% | 7% | 6% | 6% |
| All Multiple Choice | 6% | 6% | 6% | 5% | 5% |
| Free Response Question 1 | 2% | 1% | 1% | 0% | 1% |
| % Hard/Too Hard | 43% | 40% | 50% | 0% | 0% |
| % Easy/Right | 29% | 20% | 0% | 0% | 100% |
| % No Difficulty Reaction | 29% | 40% | 50% | 0% | 0% |
| Free Response Question 2 | 6% | 4% | 2% | 2% | 4% |
| % Hard/Too Hard | 58% | 46% | 40% | 33% | 33% |
| % Easy/Right | 8% | 0% | 0% | 33% | 33% |
| % No Difficulty Reaction | 33% | 54% | 60% | 33% | 33% |
| Free Response Question 3 | 7% | 6% | 4% | 3% | 4% |
| % Hard/Too Hard | 30% | 20% | 25% | 33% | 33% |
| % Easy/Right | 3% | 0% | 0% | 0% | 33% |
| % No Difficulty Reaction | 67% | 80% | 75% | 67% | 33% |
| Free Response Question 4 | 14% | 14% | 10% | 13% | 8% |
| % Hard/Too Hard | 17% | 14% | 13% | 4% | 0% |
| % Easy/Right | 3% | 0% | 0% | 0% | 0% |
| % No Difficulty Reaction | 80% | 86% | 88% | 96% | 100% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.19

### Form 3: Science
### Percent of Test Items Omitted By Coursework in Science

|  | Total | At least 1 year Biology | At least 1 year Chemistry | At least 1 year Physics |
|---|---|---|---|---|
| Number of Cases | 411 | 354 | 212 | 98 |
| Multiple Choice Science | 5% | 4% | 3% | 3% |
| Free Response Question 1 | 11% | 10% | 7% | 8% |
| % Hard/Too Hard | 72% | 73% | 66% | 55% |
| % Easy/Right | 22% | 23% | 27% | 32% |
| % No Difficulty Reaction | 6% | 4% | 7% | 14% |
| Free Response Question 2 | 11% | 10% | 7% | 7% |
| % Hard/Too Hard | 72% | 74% | 73% | 57% |
| % Easy/Right | 9% | 9% | 7% | 14% |
| % No Difficulty Reaction | 19% | 18% | 20% | 29% |
| Free Response Question 3 | 43% | 42% | 33% | 23% |
| % Hard/Too Hard | 71% | 72% | 74% | 74% |
| % Easy/Right | 10% | 11% | 11% | 9% |
| % No Difficulty Reaction | 19% | 17% | 14% | 17% |
| Free Response Question 4 | 24% | 23% | 15% | 15% |
| % Hard/Too Hard | 42% | 42% | 25% | 40% |
| % Easy/Right | 2% | 2% | 3% | 7% |
| % No Difficulty Reaction | 56% | 56% | 72% | 53% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.20

### Form 4: Science
### Percent of Test Items Omitted By Coursework in Science

| | Total | At least 1 year Biology | At least 1 year Chemistry | At least 1 year Physics |
|---|---|---|---|---|
| Number of Cases | 367 | 324 | 192 | 105 |
| Multiple Choice Science | 5% | 5% | 6% | 4% |
| Free Response Question 1 | 19% | 18% | 13% | 12% |
| % Hard/Too Hard | 73% | 76% | 76% | 69% |
| % Easy/Right | 11% | 10% | 4% | 8% |
| % No Difficulty Reaction | 16% | 14% | 20% | 23% |
| Free Response Question 2 | 37% | 33% | 24% | 19% |
| % Hard/Too Hard | 69% | 65% | 63% | 45% |
| % Easy/Right | 15% | 19% | 17% | 30% |
| % No Difficulty Reaction | 16% | 16% | 20% | 25% |
| Free Response Question 3 | 18% | 16% | 13% | 10% |
| % Hard/Too Hard | 60% | 58% | 54% | 60% |
| % Easy/Right | 3% | 4% | 8% | 0% |
| % No Difficulty Reaction | 37% | 38% | 38% | 40% |
| Free Response Question 4 | 59% | 56% | 36% | 38% |
| % Hard/Too Hard | 71% | 71% | 60% | 65% |
| % Easy/Right | 5% | 4% | 6% | 5% |
| % No Difficulty Reaction | 24% | 25% | 34% | 30% |

Note: Percentages of student difficulty evaluations are based on the students omitting the item.

## Table 3.10.22
## Correlations of Free Response Scale Scores with Ethnicity and Gender Before and After Controlling for Total Multiple Choice Score

| | Black + Hispanic vs. White + Asian | | Male vs. Female | |
|---|---|---|---|---|
| | Original Correl | Partial Correl | Original Correl | Partial Correl |
| Form 1 (Mathematics): | | | | |
| Question 1 | -.26 | -.05 | .04 | .00 |
| Question 2 | -.29 | -.13 | .08 | .06 |
| Question 3 | -.29 | -.13 | .00 | -.04 |
| Question 4 | -.16 | -.06 | .06 | .04 |
| Form 2 (Mathematics): | | | | |
| Question 1 | -.21 | -.17 | .07 | .01 |
| Question 2 | -.19 | -.15 | .03 | -.03 |
| Question 3 | -.19 | -.14 | .10 | .06 |
| Question 4 | -.18 | -.14 | .08 | .04 |
| Form 3 (Science): | | | | |
| Question 1 | -.24 | -.15 | .20 | .13 |
| Question 2 | -.12 | .07 | .22 | .12 |
| Question 3 | -.14 | .00 | .20 | .11 |
| Question 4 | -.21 | -.08 | .01 | -.11 |
| Form 4 (Science): | | | | |
| Question 1 | -.23 | -.13 | .10 | .04 |
| Question 2 | -.17 | -.07 | .18 | .13 |
| Question 3 | -.24 | -.16 | .09 | .05 |
| Question 4 | -.10 | .02 | -.09 | -.18 |
| Form 5 (Reading Comprehension): | | | | |
| Question 1 | -.24 | -.16 | .06 | .10 |
| Question 2 | -.11 | -.00 | -.14 | -.11 |
| Question 3 | -.10 | -.01 | -.17 | -.15 |
| Question 4 | -.09 | .02 | .02 | .01 |

Note: Positive correlations indicate items on which minorities and males do better; negative correlations favor non-minorities and females. Correlation coefficients of approximately .10 or greater are statistically significant for these samples of about 300 to 400 students per item.

## Table 3.10.23

## Counts of Imputed Scores

| | Related Course | # Imputes | # with High M.C. | # Taking Course |
|---|---|---|---|---|
| Form 1: | | | | |
| Question 1 | Geometry | 9 | 0 | 1 |
| Question 2 | Algebra | 13 | 1 | 8 |
| Question 3 | Algebra | 20 | 1 | 13 |
| Question 4 | Geometry | 70 | 17 | 27 |
| Form 2: | | | | |
| Question 1 | Algebra | 3 | 1 | 2 |
| Question 2 | Geometry | 14 | 4 | 2 |
| Question 3 | N.A. | 8 | 1 | N.A. |
| Question 4 | Algebra | 10 | 4 | 7 |
| Form 3: | | | | |
| Question 1 | N.A. | 1 | 0 | N.A. |
| Question 2 | N.A. | 34 | 7 | N.A. |
| Question 3 | Physics | 125 | 48 | 17 |
| Question 4 | Chemistry | 41 | 10 | 8 |
| Form 4: | | | | |
| Question 1 | N.A. | 51 | 16 | N.A. |
| Question 2 | Chemistry | 93 | 23 | 28 |
| Question 3 | Biology | 39 | 9 | 30 |
| Question 4 | Chemistry | 152 | 59 | 41 |
| Form 5: | | | | |
| Question 1 | English | 11 | 2 | N.A. |
| Question 2 | English | 5 | 0 | N.A. |
| Question 3 | English | 4 | 0 | N.A. |
| Question 4 | English | 3 | 0 | N.A. |

(N.A. for some items indicates that the material would not necessarily have been taught in a particular course, or that students would have been exposed to it in a pre-high school course such as general science.)

## Table 3.10.24
## Statistics on Reader Agreement: Form 1

### Question 1: Square Grid Problem

Subscores:

| | |
|---|---|
| area | 94% |
| 1/2perim | 81% |
| 1/2 area | 85% |
| exp-math | 75% |
| exp-clar | 38% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 5 | | 1 | | | |
| d | 2 | 1 | 20 | 1 | | | |
| r | 3 | | | 15 | 1 | | |
| 2 | 4 | | | | 1 | 1 | |
| | 5 | | | | | 3 | |
| | na | | | | | | 3 |

Scaled score: 90% agreement
8% off by 1 score point

### Question 2: Balance Beam

Subscores:

| | |
|---|---|
| part a | 91% |
| part b | 98% |
| part c | 81% |
| part d | 74% |

| | | Reader 1 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | na |
| R | 1 | 3 | 2 | 1 | | | | | |
| d | 2 | | 12 | 1 | | | | | |
| r | 3 | | | 13 | 3 | | | | |
| 2 | 4 | | | | 3 | 1 | | | |
| | 5 | | | | | 6 | | | |
| | 6 | | | | | | 5 | | |
| | 7 | | | | | | | | |
| | na | | | | | | | | 3 |

Scaled score: 85% agreement
13% off by 1 score point

### Question 3: Number System

Subscores:

| | |
|---|---|
| part a | 91% |
| part b | 97% |
| part c | 74% |
| part d | 66% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 5 | | 1 | | | 1 |
| d | 2 | | | | | | |
| r | 3 | | | 3 | 1 | | |
| 2 | 4 | | | 1 | 15 | 1 | |
| | 5 | | | | 1 | 1 | |
| | na | 1 | | | | | 4 |

Scaled score: 80% agreement
11% off by 1 score point

### Question 4: Equilateral and Isosceles Triangles

Subscores:

| | |
|---|---|
| a:math | 62% |
| a:clear | 21% |
| b:math | 69% |
| b:clear | 25% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 21 | 3 | | 1 | | 3 |
| d | 2 | 2 | 2 | | 1 | | 1 |
| r | 3 | 3 | | | 1 | | |
| 2 | 4 | | | | | | |
| | 5 | | | | | | |
| | na | 1 | 1 | | | | 12 |

Scaled score: 67% agreement
12% off by 1 score point

# Table 3.10.25
## Statistics on Reader Agreement: Form 2

### Question 1: Train Schedule

| Subscores: | |
|---|---|
| part a | 98% |
| part b | 98% |
| part c | 100% |
| part d | 74% |

| | | Reader 1 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | na |
| R | 1 | 1 | | | | | | |
| d | 2 | | 12 | 1 | | | | |
| r | 3 | | 1 | 16 | | 3 | | |
| 2 | 4 | | | | 16 | | | |
| | 5 | | | | | 3 | | |
| | 6 | | | | | 2 | | |
| | na | | | | | | 3 | |

Scaled score: 91% agreement
3% off by 1 score point

### Question 2: Area of Figure Made of Rectangles

| Subscores: | |
|---|---|
| part a | 95% |
| part b | 95% |
| part c | 89% |
| used b | 52% |
| part d | 80% |
| d: units | 55% |

| | | Reader1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 7 | 1 | | | | |
| d | 2 | 1 | 8 | | | | 1 |
| r | 3 | | 2 | 8 | | | |
| 2 | 4 | | | 1 | 1 | 1 | |
| | 5 | | | | | 8 | |
| | na | | 1 | | | | 4 |

Scaled score: 82% agreement
14% off by 1 score point

### Question 3: Dice Problem

| Subscores: | |
|---|---|
| part a | 91% |
| part b | 91% |
| part c | 83% |
| expl | 68% |

| | | Reader 1 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | na |
| R | 1 | 9 | | | | | | |
| d | 2 | | 22 | | | | | |
| r | 3 | | | 13 | | | | |
| 2 | 4 | | | | | | | |
| | 5 | | | | | | | |
| | 6 | | | | | | 2 | |
| | na | | | | | | | 7 |

Scaled score:100% agreement
0% off by 1 score point

### Question 4: Car Stopping Distance

| Subscores: | |
|---|---|
| part a | 93% |
| part b | 91% |
| part c | 72% |
| part d | 72% |

| | | Reader 1 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | na |
| R | 1 | 5 | 1 | | | 2 | | |
| d | 2 | | 17 | | | 3 | | |
| r | 3 | | 1 | | | | | |
| 2 | 4 | | | | 3 | | | |
| | 5 | 1 | 2 | | 1 | 5 | | |
| | 6 | 1 | | | | 1 | 6 | |
| | na | | | | | | | 8 |

Scaled score: 77% agreement
7% off by 1 score point

## Table 3.10.26
## Statistics on Reader Agreement: Form 3

### Question 1A: Figural Response (Pulling Rock)

Subscores:

| | |
|---|---|
| pullrock | 100% |

| | Reader 1 | | | |
|---|---|---|---|---|
| Rdr 2 | 1 | 2 | 3 | na |
| 1 | 2 | | | |
| 2 | | 1 | | |
| 3 | | | 44 | |
| na | | | | 3 |

Scaled score:100% agreement
0% off by 1 score point

### Question 1B: Figural Response (Blood Flow)

Subscores:

| | |
|---|---|
| venacava | 98% |
| aorta | 92% |
| left atr | 96% |
| rt vent | 84% |
| gut | 88% |

| | Reader 1 | | | | | | |
|---|---|---|---|---|---|---|---|
| Rdr 2 | 1 | 2 | 3 | 4 | 5 | 6 | na |
| 1 | 5 | 1 | | | | | 1 |
| 2 | | 2 | 2 | | | | |
| 3 | 1 | 1 | 9 | | | | |
| 4 | | | | 6 | 1 | | |
| 5 | | | | 2 | 6 | | |
| 6 | | | | | | 5 | |
| na | | | | | | | 8 |

Scaled score: 82% agreement
14% off by 1 score point

### Question 1C: Figural Response (Wind Direction)

Subscores:

| | |
|---|---|
| wind dir | 72% |

| | Reader 1 | | |
|---|---|---|---|
| Rdr 2 | 1 | 2 | na |
| 1 | 40 | 1 | 3 |
| 2 | 1 | 1 | |
| na | | | 4 |

Scaled score: 90% agreement
4% off by 1 score point

### Question 1D: Figural Response (Rotating ball)

Subscores:

| | |
|---|---|
| ball | 72% |

| | Reader 1 | | |
|---|---|---|---|
| Rdr 2 | 1 | 2 | na |
| 1 | 23 | 3 | |
| 2 | 1 | 16 | |
| na | 2 | | 5 |

Scaled score: 88% agreement
8% off by 1 score point

## Table 3.10.26 (continued)
## Statistics on Reader Agreement: Form 3 (continued)

### Question 1E: Figural Response (Radioactive Decay)

Subscores:

| | |
|---|---|
| decay | 84% |

| | Reader 1 | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | na |
| R 1 | 20 | 2 | | 2 |
| d 2 | | 4 | 2 | |
| r 3 | | | 8 | |
| 2 na | | 1 | | 11 |

Scaled score: 86% agreement
8% off by 1 score point

### Question 2: Eclipses

Subscores:

| | |
|---|---|
| part a | 87% |
| part b | 84% |
| part c | 78% |

| | Reader 1 | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | na |
| R 1 | 15 | | | | | |
| d 2 | 1 | 10 | | 1 | | |
| r 3 | | | 5 | 1 | 2 | |
| 2 4 | | | | | | |
| 5 | | | | | 6 | |
| na | | | | | | 4 |

Scaled score: 89% agreement
4% off by 1 score point

### Question 3: Colliding Railroad Cars

Subscores:

| | |
|---|---|
| speed | 78% |
| eqns | 89% |
| vars def | 78% |

| | Reader 1 | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | na |
| R 1 | 14 | 1 | | | 4 |
| d 2 | 1 | 3 | | | |
| r 3 | | | 1 | | |
| 2 4 | | | | | |
| na | | | | | 13 |

Scaled score: 84% agreement
5% off by 1 score point

### Question 4: Heating Curve

Subscores:

| | |
|---|---|
| sect. a | 73% |
| sect. b | 68% |
| sect. c | 77% |

| | Reader 1 | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | na |
| R 1 | 2 | 1 | | | |
| d 2 | 1 | 15 | | 2 | |
| r 3 | 2 | | 5 | 4 | |
| 2 4 | | 2 | 1 | 2 | |
| na | | | | | 7 |

Scaled score: 70% agreement
16% off by 1 score point

## Table 3.10.27
## Statistics on Reader Agreement: Form 4

### Question 1: Position of Earth on 6/21 and 12/21

Subscores:

| | |
|---|---|
| drawing | 95% |
| axis | 38% |
| orbit | 55% |
| distance | 58% |
| expl | 53% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 14 | | | | | |
| d | 2 | 1 | 1 | | 1 | | |
| r | 3 | 3 | 1 | 3 | 2 | | |
| 2 | 4 | | 2 | | 4 | | |
| | 5 | | | | | 1 | |
| | na | 1 | | | | | 6 |

Scaled score: 73% agreement
10% off by 1 score point

### Question 2: Nuclear vs. Fossil Fuels

Subscores:

| | |
|---|---|
| any resp | 77% |
| # adv | 15% |
| #reasons | 21% |
| #incor. | 34% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | | | | 1 | | 2 |
| d | 2 | 3 | 1 | | | | |
| r | 3 | | 1 | | 3 | | |
| 2 | 4 | | 2 | 1 | 2 | 6 | |
| | 5 | | 2 | 2 | 1 | 7 | |
| | na | 3 | | | | | 16 |

Scaled score: 49% agreement
28% off by 1 score point

### Question 3: Rabbit/Wolf Populations

Subscores:

| | |
|---|---|
| any draw | 88% |
| phase | 56% |
| lead/lag | 53% |
| heights | 44% |
| any expl | 65% |
| exp ampl | 35% |
| exp lag | 32% |
| r to w | 29% |
| w to r | 44% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 8 | | 3 | | | |
| d | 2 | 1 | 1 | 2 | | | |
| r | 3 | | | 3 | 3 | 1 | |
| 2 | 4 | | 1 | | 5 | | |
| | 5 | | | | | | |
| | na | | | | | | 6 |

Scaled score: 68% agreement
18% off by 1 score point

### Question 4: Stoichiometry

Subscores:

| | |
|---|---|
| equation | 93% |
| m.w. ch4 | 28% |
| #mol ch4 | 30% |
| mole rat | 28% |
| #mol h20 | 25% |
| m.w. h20 | 28% |
| mass h20 | 20% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 4 | | | | | 1 |
| d | 2 | | 6 | | | 1 | |
| r | 3 | | 1 | 1 | | | |
| 2 | 4 | | | | 4 | | |
| | 5 | | 1 | | | 2 | |
| | na | | | | | | 19 |

Scaled score: 90% agreement
3% off by 1 score point

## Table 3.10.28
## Statistics on Reader Agreement: Form 5

### Question 1: Forget That House?

Subscores:

| | |
|---|---|
| outline | 55% |
| labels | 74% |
| roof | 89% |
| bldg ht | 81% |
| tower ht | 79% |
| tower w | 92% |
| bldg w | 81% |
| bldg d | 92% |
| embell | 49% |

| | | Reader 1 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | na |
| R | 1 | | | | | | | 1 |
| d | 2 | | 19 | 1 | | 3 | 4 | |
| r | 3 | | | 1 | | | | |
| 2 | 4 | | | | | | | |
| | 5 | | | | | 11 | | |
| | 6 | | 1 | | | 3 | 7 | |
| | na | | 1 | | | | | 1 |

Scaled score: 74% agreement
8% off by 1 score point

### Question 2: Producing Rice for the Table

Subscores:

| | |
|---|---|
| any resp | 100% |
| main pts | 96% |
| step 5 | 89% |
| adl stps | 93% |
| relevant | 72% |
| order | 100% |
| number | 98% |

| | | Reader 1 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | na |
| R | 1 | 1 | | | | | | | |
| d | 2 | | 5 | | | | | | |
| r | 3 | | | 12 | | | | 1 | |
| 2 | 4 | | | 1 | 9 | | | | |
| | 5 | | | | | 15 | | | |
| | 6 | | | | | | 2 | 1 | |
| | 7 | | | | | | | | |
| | na | | | | | | | | 7 |

Scaled score: 94% agreement
4% off by 1 score point

### Question 3: Letter of Recommendation

Subscores:

| | |
|---|---|
| decision | 100% |
| reasons | 100% |
| # acc | 44% |
| # unacc | 51% |
| hidden m | 82% |
| expl | 74% |
| clarity | 82% |

| | | Reader 1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | na |
| R | 1 | | | | | | | | | |
| d | 2 | | 8 | | 1 | | 3 | | | |
| r | 3 | | 1 | | | | | | | |
| 2 | 4 | | 4 | | | | | 1 | | |
| | 5 | | | | | | | 1 | | |
| | 6 | | 1 | | 1 | | 2 | 1 | | |
| | 7 | | 1 | | | 4 | | 6 | 2 | |
| | 8 | | | | | | | | | |
| | na | | | | | | | | | 2 |

Scaled score: 46% agreement
10% off by 1 score point

### Question 4: Health-GLO

Subscores:

| | |
|---|---|
| quest 1 | 87% |
| quest 2 | 77% |
| quest 3 | 81% |
| reasos | 81% |

| | | Reader 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | na |
| R | 1 | 1 | | | | | |
| d | 2 | | 7 | 1 | | | |
| r | 3 | | 4 | 26 | 2 | | |
| 2 | 4 | | | | 6 | | |
| | 5 | | | | | 1 | |
| | na | | | | | | 5 |

Scaled score: 87% agreement
13% off by 1 score point

## Sample of Free Response Items

This section gives examples of the free response questions used in the field test: one in each of the three subject areas, mathematics, science, and reading comprehension. In the test booklets, each item was followed by the short series of student-reaction questions that is shown in this appendix only with the first item. After each question is the scoring guide that defines the analytic scores used to classify features of the students' responses. Note that these analytic scores are categorical, and are not intended to be monotonically increasing. The following page has the specifications that were used to convert the multiple analytic scores into a scale.

## Scoring Guide
## Form II, Question 2: Area of Figure Made of Rectangles

6 analytic scores:

--- **A. Area of Decomposed Figure**

0 = No attempt
1 = Off topic
2 = Either method of solving problem is incorrect or no method is shown
3 = Uses addition instead of multiplication to solve problem
4 = Uses correct method to solve problem but makes arithmetic error
5 = 26 (correct)

--- **B. Draw Lines to Decompose Figure into Rectangles**
0 = No attempt
1 = Off topic
2 = Incorrect
3 = Correct decomposition into rectangles (may include extraneous lines filling out full rectangle)

--- **C. Area of Figure in B**
0 = No attempt
1 = Off topic
2 = Either method of solving problem is incorrect or no method is shown
3 = Calculates perimeter (32) instead of area
Method for calculating area correct, but:
4 = Addition of areas of rectangles done incorrectly
5 = Areas of individual rectangle(s) are determined incorrectly
6 = 30 (correct)

--- **If Attempted to do (C) - used (B) to do (C)?**
1 = yes
2 = no
3 = can't tell

--- **D. Area of Shaded Region**
0 = No attempt
1 = Off topic
2 = Either method of solving problem is incorrect or no method is shown
3 = Made some attempt related to the problem but did not get far
4 = Attempted to break figure into rectangles but did not do anything else
5 = Correctly calculates area of large figure but ignores that small figure is not shaded
6 = Correctly calculates area of large figure (44 or $44a^2$) but incorrectly calculates area of small figure (correct area is 14 or $14a^2$)
7 = 30 (correct)
8 = Correctly cauculates area of small figure but ignores large figure
9 = Uses correct method but makes arithmetic error

--- **Units**
1 = Incorrect (either no unit or, for example, a)
2 = Incorrect ("square units" or "a square units")
3 = Correct ($a^2$ or $a^2$ square units)

## Specifications for Scale Score
## Form II, Question 2: Area of Figure Made of Rectangles

6 scores:

Part A: Area of Figure
Part B: Draw lines to decompose figure into rectangles
Part C: Area of figure in B
Used (B) to do (C)
Part D: Area of shaded region
Part D: used correct units

**Scaled Score:**

1 = any part attempted (≥ code 2) but nothing correct
2 = part A essentially correct (code 4 or 5)
3 = part B correct (code 3) and part C essentially correct (4-6)
4 = score 3, plus part D code 5, 6 or 8
5 = score 3, plus part D code 7 or 9
x = inconsistent response: complex figure essentially correct
(part D ≥ 5) but parts A or B incorrect

## Scoring Guide
## Form IV, Question 3: Rabbit/Wolf Populations

9 scores:

--- **Any Drawing?**

0 = No drawing
1 = Incomprehensible drawing
2 = Comprehensible drawing, even if incorrect

Three features of drawing--score these only if there is a comprehensible drawing:

--- **Curves Are in Phase?**

0 = The wolf curve is out of phase with the rabbit curve
1 = The wolf curve mimics the function of the rabbit curve

--- **Wolf curve leads or lags?**

0 = The wolf curve leads the rabbit curve
1 = The wolf curve changes direction at the same time as the rabbit curve
2 = The wolf curve lags the rabbit curve

--- **Relative heights of curves**

0 = Wolf curve is higher than the rabbit curve
1 = Same height
2 = Wolf curve is lower than the rabbit curve

--- **Explanation:** (score even if there is no drawing)

0 = No explanation
1 = Irrelevant or incomprehensible explanation
2 = Comprehensible explanation, even if incorrect

Four features of explanation--score these only if there is a comprehensible explanation:

--- **A. The lower amplitude of the wolf curve**

0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

--- **B. The wolf curve lags behind the rabbit curve.**

0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

--- **C. More rabbits makes possible more wolves and fewer rabbits results in fewer wolves.**

0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

--- **D. Fewer wolves makes possible more rabbits.**

0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

## Specifications for Scale Score
## Form IV, Question 3: Rabbit/Wolf Populations

9 analytic scores:

Any drawing?
Curves are in phase?
Wolf curve leads or lags?
Relative heights of curves
Any explanation?
Explained lower amplitude of the wolf curve
Explained lag of wolf curve
Rabbit numbers → wolf numbers
Wolf numbers → rabbit numbers

Scaled Score:

1 = Drawing present (1 or 2), and/or explanation (1-3) but not correct
2 = One correct feature in drawing (Phase, lag, height)
3 = Two correct feeatures
4 = Three correct features

Add one point if answer has one or more correct explanations.
Subtract one point if answer has one or more incorrect explanations
(but score may not be less than 1)

## Scoring Guide
## Form V, Question 2: Producing Rice for the Table

6 analytic scores:

--- **A. Any Response?**

0 = No response
1 = Response is incomprehensible or deliberately inappropriate
2 = "I don't know how to do this" or "I can't do this"
3 = Comprehensible response, even if incorrect

Five features--score these only if there is a comprehensible response:
(Otherwise leave blank)

--- **B. Main Points** (excluding optional steps):

- ★ drying
- ★ coarse cleaning
- ★ milling to remove the husk
- ★ milling to remove the bran

0 = no main points mentioned
1 = one main point mentioned
2 = two main points mentioned
3 = three main points mentioned
4 = all four main points mentioned

--- **C. Step #5**: optional processing such as adding vitamins

0 = Not included
1 = Included without mention that it is optional
2 = Recognized as optional

--- **D. Any additional steps**

0 = none
1 = one or two others (such as farming steps, storage, etc.)
2 = three or more others

--- **E. Relevance**

0 = Some topics included are not relevant (seeding, harvesting, etc.)
1 = All topics included are relevant

--- **F. Order**

0 = Some topics are out of order
1 = Topics are in proper order

## Specifications for Scale Score
## Form V, Question 2: Producing Rice for the Table

7 analytic scores:

Any Response?
Main Points Present
Optional Step #5
Any additional steps
Relevance
Order
Numbering

**Scaled Score:**

1 = Question attempted, but no main points listed (A=2 or 3; B=0)
2 = At least one main point (B≥1), or more than one point with something out of order (F=0); may have addl/irrelevant steps
3 = At least two main points, order correct (B≥2, F=1); may have additional or irrelevant steps
4 = At least three main points, order correct (B≥3, F=1); no more than two additional steps (D≤1); may have irrelevant information
5 = Four main points, order correct (B=4,F=1), no more than two additional steps (D≤1); may have irrelevant information
6 = Four main points, order correct, nothing irrelevant, no additional steps (B=4,F=1,E=1,D=0)
7 = Same as score 6, plus Step #5 recognized as optional (C=2)

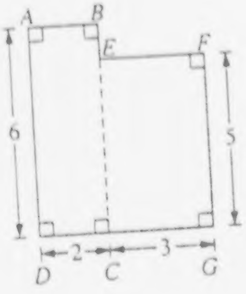
Question 2

To find the area of a figure made up of two or more rectangles, we can find the area of each rectangle and add the areas together. For example:

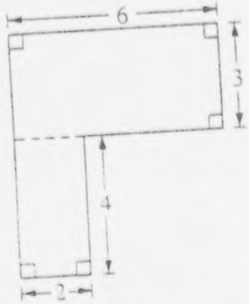
Area of rectangle $ABCD$ = 12 square units

Area of rectangle $CEFG$ = 15 square units

Total area = 27 square units

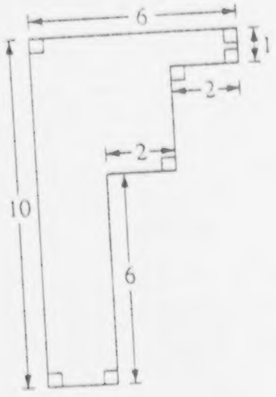


(A) What is the area of the figure below?



Answer: ______________________________

(B) Draw lines in the figure below to show that it is made up of several rectangles.



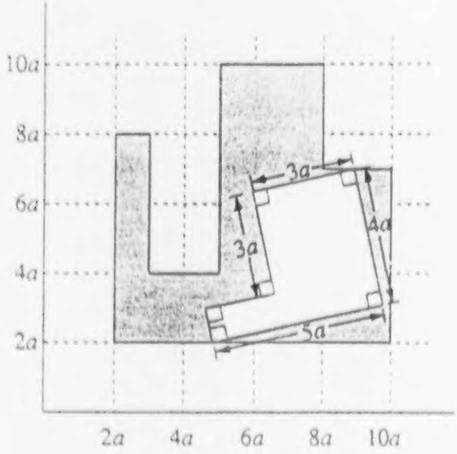
(C) What is the area of the figure in B?

Answer: ______________________________

GO ON TO THE NEXT PAGE.

(D) Determine, in terms of $a$, the area of the shaded region below

Answer: ____________________



For each of the following, circle the phrase that best describes how you did on this question.

1. How hard was the question?

   (A) Too easy　(B) Easy　(C) About right　(D) Hard　(E) Too hard

2. Did you have enough time to answer the question?

   (A) Not enough time at all
   (B) Could have used a little more time
   (C) About the right amount of time
   (D) A little too much time
   (E) Way too much time

3. Did you understand the question?

   (A) It was very clear.
   (B) It was clear enough.
   (C) It was a little confusing.
   (D) It was very confusing.

4. Did you give the best answer you could?

   (A) Yes
   (B) No: Why not? ____________________

STOP

## Scoring Guide
## Form II, Question 2: Area of Figure Made of Rectangles

6 analytic scores:

--- A. Area of Decomposed Figure

0 = No attempt
1 = Off topic
2 = Either method of solving problem is incorrect or no method is shown
3 = Uses addition instead of multiplication to solve problem
4 = Uses correct method to solve problem but makes arithmetic error
5 = 26 (correct)

--- B. Draw Lines to Decompose Figure into Rectangles
0 = No attempt
1 = Off topic
2 = Incorrect
3 = Correct decomposition into rectangles (may include extraneous lines filling out full rectangle)

--- C. Area of Figure in B
0 = No attempt
1 = Off topic
2 = Either method of solving problem is incorrect or no method is shown
3 = Calculates perimeter (32) instead of area
Method for calculating area correct, but:
4 = Addition of areas of rectangles done incorrectly
5 = Areas of individual rectangle(s) are determined incorrectly
6 = 30 (correct)

--- If Attempted to do (C) - used (B) to do (C)?
1 = yes
2 = no
3 = can't tell

--- D. Area of Shaded Region
0 = No attempt
1 = Off topic
2 = Either method of solving problem is incorrect or no method is shown
3 = Made some attempt related to the problem but did not get far
4 = Attempted to break figure into rectangles but did not do anything else
5 = Correctly calculates area of large figure but ignores that small figure is not shaded
6 = Correctly calculates area of large figure (44 or $44a^2$) but incorrectly calculates area of small figure (correct area is 14 or $14a^2$)
7 = 30 (correct)
8 = Correctly cauculates area of small figure but ignores large figure
9 = Uses correct method but makes arithmetic error

--- Units
1 = Incorrect (either no unit or, for example, a)
2 = Incorrect ("square units" or "a square units")
3 = Correct ($a^2$ or $a^2$ square units)

## Specifications for Scale Score
## Form II, Question 2: Area of Figure Made of Rectangles

6 scores:

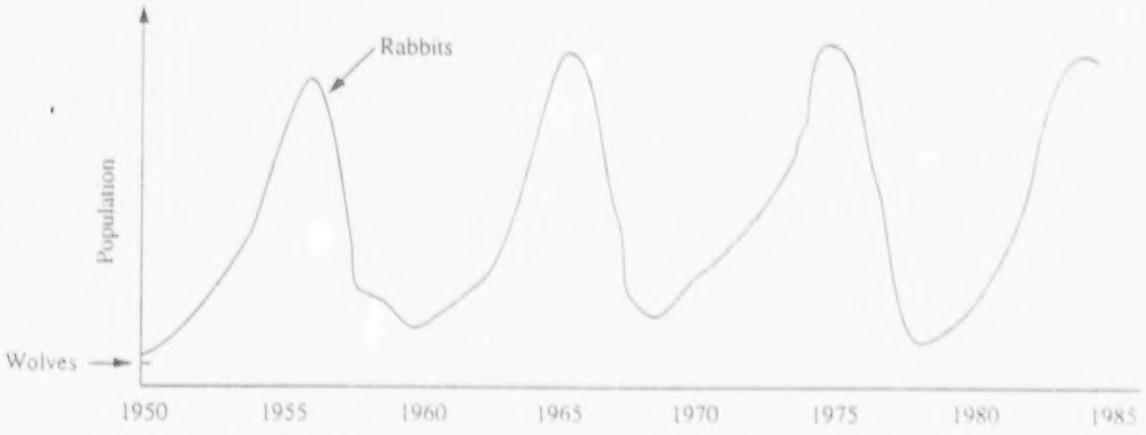
Part A: Area of Figure
Part B: Draw lines to decompose figure into rectangles
Part C: Area of figure in B
Used (B) to do (C)
Part D: Area of shaded region
Part D: used correct units

Scaled Score:

1 = any part attempted (≥ code 2) but nothing correct
2 = part A essentially correct (code 4 or 5)
3 = part B correct (code 3) and part C essentially correct (4-6)
4 = score 3, plus part D code 5, 6 or 8
5 = score 3, plus part D code 7 or 9
x = inconsistent response: complex figure essentially correct
(part D ≥ 5) but parts A or B incorrect

Question 3

A particular species of rabbit is infected with a rabbit-specific virus that is only active when the population of rabbits reaches a specific density. The virus characteristically decimates the population at fairly regular intervals, as shown on the graph below. The rabbits share their ecosystem on an isolated island with a species of wolf for which the rabbit is the predominant prey. On the same graph below, draw a curve that might reasonably represent the population of wolves over the same time period, starting with the population point given. On the lines below the graph, briefly explain the shape and location of the curve you draw.



GO ON TO THE NEXT PAGE.

3C0

## Scoring Guide
## Form IV, Question 3: Rabbit/Wolf Populations

9 scores:

--- Any Drawing?
0 = No drawing
1 = Incomprehensible drawing
2 = Comprehensible drawing, even if incorrect

Three features of drawing--score these only if there is a comprehensible drawing:

--- Curves Are in Phase?
0 = The wolf curve is out of phase with the rabbit curve
1 = The wolf curve mimics the function of the rabbit curve

--- Wolf curve leads or lags?
0 = The wolf curve leads the rabbit curve
1 = The wolf curve changes direction at the same time as the rabbit curve
2 = The wolf curve lags the rabbit curve

--- Relative heights of curves
0 = Wolf curve is higher than the rabbit curve
1 = Same height
2 = Wolf curve is lower than the rabbit curve

--- Explanation: (score even if there is no drawing)
0 = No explanation
1 = Irrelevant or incomprehensible explanation
2 = Comprehensible explanation, even if incorrect

Four features of explanation--score these only if there is a comprehensible explanation:

--- A. The lower amplitude of the wolf curve
0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

--- B. The wolf curve lags behind the rabbit curve.
0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

--- C. More rabbits makes possible more wolves and fewer rabbits results in fewer wolves.
0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

--- D. Fewer wolves makes possible more rabbits.
0 = not mentioned
1 = mentioned but incorrectly
2 = explained correctly

## Specifications for Scale Score
## Form IV, Question 3: Rabbit/Wolf Populations

9 analytic scores:

Any drawing?
Curves are in phase?
Wolf curve leads or lags?
Relative heights of curves
Any explanation?
Explained lower amplitude of the wolf curve
Explained lag of wolf curve
Rabbit numbers → wolf numbers
Wolf numbers → rabbit numbers

Scaled Score:

1 = Drawing present (1 or 2), and/or explanation (1-3) but not correct
2 = One correct feature in drawing (Phase, lag, height)
3 = Two correct feeatures
4 = Three correct features

Add one point if answer has one or more correct explanations.
Subtract one point if answer has one or more incorrect explanations
(but score may not be less than 1)

Question 2. Read this passage carefully.

## PRODUCING RICE FOR THE TABLE

Harvesting rice in the United States is a mechanized process. Tractors, threshers, and loading machinery all lighten the work of rice farmers. Even airplanes may be used in seeding, fertilizing, and controlling pests.

When it is harvested, the kernel of rice has a hull or husk. This rice is known as paddy or rough rice. Before this rice is processed, it is mechanically dried to bring the moisture content down to about 13 to 14 percent, which is low enough to prevent the grain from spoiling in storage. During commercial drying, a coarse cleaning removes foreign materials. After drying, the rice is either stored for use as seed or animal feed or taken to the milling and processing plant.

At the processing plant, the first step in rice milling is removing the husk from each grain of rice. After this has been done, the rice is called *brown rice*. Only a small amount is marketed in this form in the United States. Brown rice is the most nutritious of the rice forms; it has a nutlike flavor and a chewy texture.

The form of rice that is most familiar to the rice eaters in the United States is *white milled rice*, from which several layers of bran have been removed. The process of removing the bran to reveal the familiar white rice kernel takes several steps and two or three types of machines. After the removal of the bran, the rice is ready for such optional processing steps as adding vitamins to the grain to produce *enriched rice*.

List briefly the steps that are taken after harvesting to produce the white rice usually eaten in the United States. Number each step.

GO ON TO THE NEXT PAGE.

## Scoring Guide
### Form V, Question 2: Producing Rice for the Table

6 analytic scores:

--- A. Any Response?

0 = No response
1 = Response is incomprehensible or deliberately inappropriate
2 = "I don't know how to do this" or "I can't do this"
3 = Comprehensible response, even if incorrect

Five features--score these only if there is a comprehensible response:
(Otherwise leave blank)

--- B. Main Points (excluding optional steps):

drying
coarse cleaning
milling to remove the husk
milling to remove the bran

0 = no main points mentioned
1 = one main point mentioned
2 = two main points mentioned
3 = three main points mentioned
4 = all four main points mentioned

--- C. Step #5: optional processing such as adding vitamins

0 = Not included
1 = Included without mention that it is optional
2 = Recognized as optional

--- D. Any additional steps

0 = none
1 = one or two others (such as farming steps, storage, etc.)
2 = three or more others

--- E. Relevance

0 = Some topics included are not relevant (seeding, harvesting, etc.)
1 = All topics included are relevant

--- F. Order

0 = Some topics are out of order
1 = Topics are in proper order

## Specifications for Scale Score
## Form V, Question 2: Producing Rice for the Table

7 analytic scores:

Any Response?
Main Points Present
Optional Step #5
Any additional steps
Relevance
Order
Numbering

Scaled Score:

1 = Question attempted, but no main points listed (A=2 or 3; B=0)
2 = At least one main point (B≥1), or more than one point with something out of order (F=0); may have addl/irrelevant steps
3 = At least two main points, order correct (B≥2, F=1); may have additional or irrelevant steps
4 = At least three main points, order correct (B≥3, F=1); no more than two additional steps (D≤1); may have irrelevant information
5 = Four main points, order correct (B=4,F=1), no more than two additional steps (D≤1); may have irrelevant information
6 = Four main points, order correct, nothing irrelevant, no additional steps (B=4,F=1,E=1,D=0)
7 = Same as score 6, plus Step #5 recognized as optional (C=2)

# APPENDIX D

## Summary of Cognitive Interviews

## NELS:88 Second Follow-Up Field Test
## Summary of Cognitive Interviews

This appendix contains a summary of the cognitive interviews conducted to assess the NELS:88 Second Follow-Up Field Test Not Currently In School (Dropout) Questionnaire. Information about the cognitive interview technique and its application to the NELS:88 2FU dropout component can be found in Section 5.2.

This summary contains the notes made by the NORC professional staff who conducted the interviews; some editing of the notes was performed to ensure respondent's confidentiality. The first pages describe the ten participants sociodemographically. The comments are organized by question and identified by respondent number. For each question, only those respondents who were asked the question are listed. Therefore, if the item was legitimately skipped or the section was inapplicable, the respondent's number does not appear for the question. "No finding" should be interpreted as the question having been asked, but no problem detected (synonymous with interviewer notes of "OK", "No problem", etc.). Interviewer's comments/recommendations are presented in capital letters to distinguish them from participant contributions.

# NELS:88 Second Follow-Up Field Test
## Cognitive Interview Results

### Respondent Demographics

**RESPONDENT 1** (R1)

Age: 20

Sex: Female

Race/Ethnicity: Black

Marital Status: Single

Child(ren)?: Missing

Highest Grade Completed: 10

Amount of Time Out of School: Months: 0 Years: 2

How R Recruited: Saw flyer at Employment Office

**RESPONDENT 2** (R2)

Age: 20

Sex: Female

Race/Ethnicity: Black

Marital Status: Single

Child(ren)?: No

Highest Grade Completed: 11

Amount of Time Out of School: Months: 5 Years: 3

How R Recruited: Urban League

**RESPONDENT 3** (R3)

Age: 18

Sex: Female

Race/Ethnicity: Black

Marital Status: Single

Child(ren)?: None

Highest Grade Completed: 9

Amount of Time Out of School: Months: 4 Years: 3

How R Recruited: Urban League

**RESPONDENT 4** (R4)

Age: 20

Sex: Female

Race/Ethnicity: Black

Marital Status: Single

Child(ren)?: 2

Highest Grade Completed: 11

Amount of Time Out of School: Months: 11 Years: 3

How R Recruited: Urban League

**RESPONDENT 5** (R5)

Age: 19 Sex: Female

Race/Ethnicity: Black Marital Status: Single

Child(ren)?: 1 Highest Grade R Completed: 10

Amount of Time Out of School: Months: 6 Years: 2

How R Recruited: Urban League

**RESPONDENT 6** (R6)

Age: 18 Sex: Female

Race/Ethnicity: Mexican/Puerto Rican Marital Status: Living as Married

Child(ren)?: 1 Highest Grade Completed: 10

Amount of Time Out of School: Months: 2 Years: 3

How R Recruited: Personal Network

**RESPONDENT 7** (R7)

Age: 19 Sex: Female

Race/Ethnicity: Black Marital Status: Single

Child(ren)?: 1 Highest Grade Completed: 11

Amount of Time Out of School: Months: 11 Years: 0

How R Recruited: Personal Network

**RESPONDENT 8** (R8)

Age: 18 Sex: Male

Race/Ethnicity: Black Marital Status: Single

Child(ren)?: None Highest Grade Completed: 10

Amount of Time Out of School: Months: 5 Years: 1 (Approximate)

How R Recruited: Through GED program tutor

**RESPONDENT 9** (R9)

Age: 19

Sex: Female

Race/Ethnicity: Thai

Marital Status: Single

Child(ren)?: None

Highest Grade Completed: 11

Amount of Time Out of School: Months: 4 Years: 3

How R recruited: Works in shop contacted for recruitment

**RESPONDENT 10** (R10)

Age: 19

Sex: Male

Race/Ethnicity: Black

Marital Status: Never Married

Child(ren)?: None

Highest Grade Completed: 10

Amount of Time Out of School: Months: 23 Years: 0

How R Recruited: Personal Network

## Summary of Cognitive Interview Findings

## I. Introductory Section

### 1. Questionnaire Cover/Confidentiality:

**R1:** No findings

**R2:** R had difficulty understanding the word incentive and only had a fair grasp of what the term confidentiality meant.

**R3:** R thought the "USES OF THE DATA" section sounded like a bunch of words that did not make much sense.

**R4:** No findings

**R5:** N/A

**R6:** R was able to paraphrase meaning of statement in her own words well. She wasn't quite sure what the terms "policy makers and curriculum tracking" meant. She commented that she thought most dropouts regardless of race wouldn't know either. She said that if the intent of the statement is to explain the purpose of the study, then the language should be simpler. It may be a turn off to Rs because the wording is too elite. R had basic understanding of the Confidentiality pledges. She asked what th General Education Act and Public Law were and didn't feel that they needed to be there---just more to read and can seem too imposing to Rs. In point 4, asked what "responses will be combined" meant.

**R7:** R's interpretation of the "Uses of the Data" section was that we want to know why she quit school and how the experience of dropping out was for her. Understood that confidentiality meant, "no matter what this won't get out". Understood she was assured anonymity.

**R8:** No findings

**R9:** R did not read section on uses of the data on cover page, but did read the section on confidentiality. Her summary of that section indicated that she understood its gist, though was confused by term Provisions in part 1. She reported no anxieties over any of what she read.

**R10:** R read; had no questions or problems with it.

### 2. OMB Burden Box/Purpose of Survey/Voluntary Nature:

**R1:** No findings

**R2:** No findings

**R3:** No findings

**R4:** No findings

**R5:** N/A

**R6:** R asked if her interview was going to take 3 hours. Asked about the meaning of cognitive testing. Was curious about the use of the term "burden", she felt some Rs may react to this word negatively and think it applies to them not the survey. Wanted the Self-Description box explained because she didn't understand what "copyrighted and reproduction" meant and felt most others wouldn't either. The third box she felt had been already stated on the cover. The fourth box--she believed interview not a test and she recognized her choice to skip questions. Note: although R said she felt not a test, when filling out the quex, she often would ask me if she was answering correctly because she didn't want to have wrong answers for some questions.

**R7:** Said she understood statement. On probing, she asked if the interview was going to take three hours. Also asked what "cognitive" meant; thought other

dropouts would not know what it means. Asked about the meaning of "burden estimate". Asked whether other Rs would believe the questionnaire is not a test. R said "not to themselves."
**R8:** No findings
**R9:** She read this and the three boxes below. Reported understanding, and not being worried by, any of the text.
**R10:** R read and had no problems with these.

3. General Instructions:

**R1:** R was not clear on the instructions. In the example she initially missed the skip in item D from a to c. In general, R did not read item instructions (the all cap material in parentheses) aloud, although sometimes she may have read them silently.
**R2:** R appeared to have overlooked the explanations for examples A. through C. (did not read them). She also did not read the capitalized instructions in the parentheses.
**R3:** R did not know if she was to circle her choices in items a. through c. In c. she initially thought that #1 was no. Once realizing #1 was not no she was going to write in no.
**R4:** Since there was no instruction to complete a. through c., R was not sure if she should or not.
**R5:** R stumbled over the examples; was not sure whether to answer or just follow. She also did not read the instructions in caps aloud.
**R6:** When reviewing the Instructions, R's understanding was very good. However, she did comment that she felt the examples were too biased towards whites, especially the eye color and activity items. She felt this page should be more clearly labeled as an example page. The skip instruction arrows she felt are more confusing than just having the skip instructions stand alone--the instructions should be in caps and bold print however!
**R7:** Example B R understood one answer on each line vertically. However, the "next week" concept was difficult for her -- life is too immediate to predict the future it seems. Felt arrows were too confusing.
**R8:** R did not understand that these were practice items, and answered them for himself. He did not follow the instructions for B, circling a number on only one line. He found the skip example (D) very confusing. After some thought he decided to draw a line along the dashed line leading to b.

**R9:** Understood examples A, B and C. Initially very confused by D, and reported not understanding what the term skip meant. After reading down through example R was able to understand the instructions.
**R10:** R did not realize these were instructions or examples. Thought it was 1st question, wrote "brown" by color of eyes Q b4 realizing what was going on. Perhaps the word "EXAMPLES" or a paragraph explaining that these are examples should be placed at the top of this page. Was "wowed" by 'graphic art' on p. ii. Got the feeling that if I weren't there, he would not have attempted to figure out what it meant.

## II. "Your Educational Experiences and Activities"

### Q. 6 - "When did you last attend school?"

**R1:** R initially missed the month then went back to complete it.
**R2:** R did not read the instruction aloud.
**R3:** No finding
**R4:** Initially R answered September 88 because that was the time she should have returned. She went back and changed it to June 88 which was the actual last time she was physically in school.
**R5:** R did not read the instructions aloud.
**R6:** Year recall was immediate, but not month. Did self-prompting to arrive at month. Related dropout event to the same month she found out she was pregnant with her first child. She felt that was the way most Rs would have to recall--relate dropout event to another life event that occurred at the same time. She felt we will get a lot of "dk" responses for month.
**R7:** Had no problem recalling the month and the year.
**R8:** Initially answered 08/89, but then changed answer to 12/89, writing both numbers in the place for the second number.
**R9:** No findings.
**R10:** No problem.

### Q. 7 - "What grade were you in then?"

**R1:** No finding.
**R2:** R did not read the instruction aloud.
**R3:** No finding.
**R4:** No finding.
**R5:** R did not read the instructions aloud.
**R6:** Grade recall immediate.
**R7:** At first R thought question was asking for what grade completed.
**R8:** He circled 10th grade but then wrote GED program underneath question, indicating his current studies at the center.
**R9:** No finding.
**R10:** General format issue began to be apparent here. R was inclined to circle response itself instead of number on the right.

### Q. 8 - "Did you pass that grade?"

**R1:** R left school before the end of the school year.
**R2:** R did not know if she passed since she left in the middle of the academic year, a reasonable answer.
**R3:** No finding.
**R4:** No finding.
**R5:** R did not finish that academic year.
**R6:** R asked me--"Does this mean passed your studies or going to the next grade level?" I initially responded that she should answer based on what the question meant to her. When I probed later, she said she answered assuming that it meant going to the next grade level. Also, this was the first evidence of circling the "no/yes" answer category itself instead of the 1/2. When I probed later why she circled the "no" and not the "1", she said "it's stupid to have both because you want to know whether it's yes or no".

**R7:** No problem.
**R8:** As with quite a few other questions he circled the word rather than the number.
**R9:** No finding.
**R10:** OK.

Q. 9 - **"What are the main reasons you left the last school you attended?"**

**R1:** No finding.
**R2:** R did not read the instruction aloud.
**R3:** No finding.
**R4:** No finding.
**R5:** R did not read the instructions aloud.
**R6:** R didn't have to think about answer at all. When I probed whether there were other reasons, she cited other reasons such as the gang problem at her school, teacher apathy, and her own lack of motivation (too tired). However, her main reason was getting pregnant, even though she could have continued school during her pregnancy.
**R7:** After reading the question aloud, the R said she had several reasons and thought she was only to record one. (Hadn't seen plural reasons.)
**R8:** R gave appropriate answer.
**R9:** No finding.
**R10:** R had much trouble with this open-ended Q. R puzzled over this for quite some time before I intervened. Had trouble articulating reason for leaving, as there were many. Maintained he was essentially given a choice of dropping out or being kicked out... R asked if "last school attended" refer to last high school or last "any school," such as vo-tech school or GED classes.

Q. 10 - **"Before you permanently left school, did you ever leave school for more than a month for a reason other than illness?"**

**R2:** R was not completely certain of the meaning of "permanently" and thought the question was asking if she had ever missed school anytime in a one month time period.
**R3:** R circled "No" and not #1 and missed skip to Q14.
**R4:** Prompted to read skip to Q14.
**R5:** R did not read the skip instructions aloud.
**R6:** R answered without hesitation. When I probed her understanding, she was able to confirm her answer, and suggested that the word "completely" be used in place of "permanently".
**R7:** No problem with question or skip.
**R8:** Circled the word no, then erased it and circled number. R failed to follow the skip instruction.
**R9:** After circling a number for questions 7 and 8, R circled a word for questions 10, 11 and 12. After being reminded of instruction, R circled a number for most subsequent questions, but occasionally reverted to circling a word.
**R10:** OK

**Q. 10A - "When was the very first time you left school for more than a month?"**

**R5:** R was unsure of the month she left school for more than a month.
**R8:** Realized question did not apply to him, and wrote no by the boxes.
**R9:** Said May, but wrote 04, later changing it to 05. R interpreted left school more like reduced attendance at school; she had not completely left school at this time, as she was still attending morning classes. She said she was "sort of going" to school.

**Q. 11 - "Did you ever return to school after that?"**

**R5:** R did not read the skip instructions.
**R8:** For some reason R circled yes. To save time I instructed him to move to Q14.
**R9:** No finding.

**Q. 11A - "When did you return to school?"**
**R5:** R was unsure of the month she returned to school.
**R9:** R gave a date, coinciding with coming back to do fill-in classes for just a week.

**Q. 12 - "Did you leave school again for more than a month for a reason other than illness?"**

**R5:** R read the last part of the question, "for a reason other than illness", as a second question and answered a second time.
**R9:** No finding.

**Q. 12A - "When did you leave?"**

**R5:** R did not read the instructions aloud.
**R9:** Gave a date (12/87) but was somewhat vague about when she finally left, saying that she had gradually withdrawn, going to fewer and fewer classes.

**Q. 13 - "Did you return to school again?"**

**R5:** R did not return to school, she should have skipped to Q14.
**R9:** R did not return to a high school. For Q13 and 13A the school she reported returning to was a preparatory school, attended for only about one month in late 1990, three years after leaving high school. This school was reported again in the section on alternative programs, questions 26 through 31.

**Q. 14 - "About how many school days did you miss during the 1989-90 school year?"**

**R1:** R did not seem to understand this item. She last attended school 4/89, the academic yr. 88-89, not 89-90. Nonetheless she answered the first part of the question rather than circling one as she should have.
**R2:** R left school before 89-90. It was not clear to her how to answer the question when this occurs. She did not read the second half of the question.

**R3:** R did not attend school in the 89-90 school year. She wrote in 0s.
**R4:** R did not attend school during the 89-90 school year. She entered 0s in "days" but skipped over the "Did not attend ..." item.
**R5:** R did not attend school during the 89-90 school year.
**R6:** Initially, R's reaction was that there was no way for her to remember the number of days because she would miss division to be marked absent in order to cut classes without penalty--common practice in Chicago high schools. After several probes, she finally recognized that she wasn't in school in 89-90. The R felt that most pre 89/90 dropouts will miss category unless more visible. She thought that "during the 1989-90 school year" should be the initial words in the question stem to set time frame immediately.
**R7:** R asked for clarification of "days missed". Does this include cutting division to be marked as absent? R thought this will be ambiguous for most Rs if definition is not included. Recall was tied to number of occasions she cut with friends, was at home with asthma, or went to clinic appointments. R worked very hard to count the days systematically, but felt that others would not be so diligent. When recording answer, R couldn't grasp the concept of averaging or rounding and recorded a range in the boxes.
**R8:** Indicated missing no days. Wrote no on line by Did not attend school.
**R9:** Initially missed the reference to 1989-90, and thought of 1987, but noticed the year before writing down answer.
**R10:** R did not interpret this question well at all. First entered 60. I probed him because, according to Q1, he had left school for good in May 1989. So, he entered "0". I had to point out to him that he really wanted to circle "1." R, more than a couple of times, had trouble figuring out which set of responses belonged to which Q stem. YOU KNOW WHAT MIGHT HELP HERE? SIMPLY PLACING A LINE THAT RUNS HORIZONTALLY ACROSS THE PAGE TO SEPARATE Qs.

**Q. 15 - "What is the name and address of the last school you attended?"**

**R1:** R gave the street address after reading the question aloud.
**R2:** No finding.
**R3:** After R read the question she explained that she did not know the street address, and then wrote in the name of the school, Washington, DC.
**R4:** No finding.
**R5:** R did not read the instruction aloud.
**R6:** R had no problem with this question.
**R7:** Recorded complete address, including street address.
**R8:** Wrote name of one school, then erased it and wrote in name of another school. Problem seemed to be that R was unsure when he changed school.
**R9:** Reported for the preparatory school.
**R10:** Asked again, does "last school" mean "last high school"? Why do we ask for "address" if what we want is city?

**Q. 16 - "Was this the same school you attended in the 1988-89 school year?"**

**R1:** No finding
**R2:** R did not read the instructions or all the choices aloud.
**R3::** R did not read the list. She was not enrolled in the 88-89 school year but circled #1.
**R4:** No finding.
**R5:** R did not read the instruction aloud.

**R6:** At first R had a problem because of year shift in this question to 88-89. She almost circled "was not in school in 1988-89". She recommended underlining it or larger print or putting it as question stem.
**R7:** No problem. Knew question applied to her.
**R8:** Unclear whether his answer of yes was correct.
**R9:** Initially R did not notice Not in school option, but correctly selected it on doing so.
**R10:** No finding.

**Q. 17 - "Do you plan to get a high school diploma or GED?"**

**R1:** No finding.
**R2:** R did not read the instructions aloud.
**R3:** R did not read the list and circled #2.
**R4:** No finding.
**R5:** R did not read the instructions aloud.
**R6:** Initially, R circled answer category 1 because she didn't read question nor the answer category correctly---she thought question read "have you planned" and the answer category read "no, I haven't planned". By choosing #1, she skipped to Q.20 incorrectly. Upon review, she corrected to answer category #2 at first, but then recognized that #4 is more appropriate for her circumstances, after she read answer categories 2, 3 and 4 aloud. When reading aloud, she also indicated that the last part of category #4, "or don't plan to enroll" doesn't fit with the rest of the statement in that category. When I asked whether there is wording in Q.17 that caused her to misread it, she replied no, she just wasn't in focus. She also recommends that question starts with "have you enrolled in or do you plan...".
**R7:** Didn't immediately recognize that answer category #4 fit her circumstance. R didn't read all the answer categories. To make that category clearer, R felt it should switch wording around and make "don't plan to enroll" (last part of category) a separate category.
**R8:** Wrote yes by question stem and also on the dotted lines for 03 (correct response), and a no by 05. R asked what the terms equivalency and preparatory meant.
**R9:** Correctly answered (02) but then was very confused what the vertical line connecting 02 to 04 was intended to indicate. She asked whether it meant she had to go to question 04, and even after considerable thought did not know what she should do next.
**R10:** R had trouble because he was enrolled in GED prep class, but is not currently (3 says, "I am enrolled..." and 4 says, "... have not yet enrolled"; thus, felt that none of the responses accurately covered his situation.

**Q. 18 - "When do you expect to receive a high school diploma, GED, or High School Equivalency certificate?"**

**R1:** R may have interpreted this to mean the date she takes the GED exam. Wrong skip to Q22, either Q21 or Q25. R correctly followed the skip as instructed.
**R2:** R was unable to answer the question. She plans to take the GED exam soon but was unsure as to when she would receive the GED.
**R3:** R expects to get the GED in the future but is not currently working towards it. She left the item blank since she does not know.
**R4:** R has plans to take the GED exam in June 91 which is what she entered in the

item.

**R5:** R gave the year she plans to get the CED (?). R missed the skip to Q21.

**R6:** Answered after correction made to Q.17. R really wasn't sure based on her desire to wait until after baby's born and the baby's at least several months old. She expressed that other Rs who are thinking about getting their CEDs, but haven't firmly decided yet, may also have difficulty arriving at a specific month and year, or they'll answer what they feel is an appropriate answer.

**R7:** When R filled in Q. 18., recognized skip problem in going to Q. 22 and asked for help. After several probes, R was able to tell me why Q. 22 doesn't correspond to Q. 18 and that correct skip is to Q21. Asked what she would do if no one there to assist, she said she would have skipped both Q21 and Q22. Answer for Q. 18 was first recorded as month entering (05). Corrected to August after several probes.

**R8:** Seemed to answer correctly, and correctly followed the printed skip instruction.

**R9:** She had not enrolled, and was unsure when she would, so answered by writing zeros for month and year. R missed skip and went to Q19.

**R10:** R missed skip instruction. Need arrow or some other clue here.

**Q. 19 - "When did you receive a GED or High School Equivalency certificate?"**

**R5:** R has not received a GED or certificate. She missed the skip from the preceding question.

**R9:** Question did not apply to R, and she indicated this by writing zeros for month and year.

**Q. 20 - "Since leaving school, have you enrolled in an educational institution, such as vocational or trade school, or a college?"**

**R5:** R did not read the instructions aloud.

**R6:** Answered initially when R miscoded Q.17. Found out that since she dropped out, R has attended trade school where she took secretarial courses--she didn't complete trade school however. This information would have been missed had she originally coded Q.17 correctly. Shouldn't this question be asked of all Rs?

**R9:** R had enrolled briefly in a preparatory school, but answered No.

**Q. 21 - "On the whole, do you feel that leaving school was a good decision for you?"**

**R1:** No finding

**R2:** R did not read the instruction aloud.

**R3:** No finding.

**R4:** No finding.

**R5:** R did not read the instruction aloud.

**R6:** Again, from Q.17 mis-code and recommend that all Rs get this question. R suggests taking out the phrase, "on the whole" and just start the question with "do". She also pointed out that, at first, she wasn't sure which school question meant---the high school or the trade school. She made the independent judgement that all Rs don't answer Q.20, based on her own skip error from Q.17, so Q.21 must mean the high school.

**R7:** No problem answering.

**R9:** No finding

**Q. 22 - "Please explain why you feel that way."**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R5:** R did not read the instruction aloud.
**R6:** No finding.
**R7:** R was able to explain reasons well when recording answers and when asked to paraphrase what she wrote without looking at it (I read it back to her) she was able to recall all the reasons listed. Asked about # of spaces needed to record, she said it depends. She used 5 lines. Probed about 2 basic reasons she listed for main reasons - she could give me the one she listed first in recording.
**R8:** R explained his answer to Q17 rather than Q21 due to incorrect skip instruction.

**Q. 25 - "In the last high school you attended, which of the following best describes your high school program?"**

**R1:** R did not read through the list, she appeared to choose the first one that made sense.
**R2:** R was unsure of what the question was asking. "Does that mean the subjects that we had"? She read all the choices except g. which is under the line for "Other".
**R3:** R was unsure of what was meant by high school program. She selected "General high school program" anyway.
**R4:** No finding.
**R5:** R did not read the instruction, and initially did not read the list of choices aloud.
**R6:** R recommends changing the wording from "high school program" to "the type of classes you took". She felt other Rs may confuse "your high school program" to mean the type of classes offered at the high school as opposed to the student's individual program. R coded two answers--c and 05 under d. She didn't see instruction to circle one and she didn't initially read "last high school attended" part of the question. Seems to be the same confusion about having gone to trade school after dropping out of high school, which will no doubt apply to other Rs.
**R7:** Probed whether "your H.S. program" indicated R's program or school's - when probed, she said "hers" but answered as though school's. Asked how to restate had some difficulty. The term "your H.S. Program" may be ambiguous to a number of R's.
**R8:** R appeared to answer this in terms of the training he wanted after he gets his GED, rather than training he had at school.
**R9:** R was unsure whether to report for her high school or the prep school that she had briefly attended recently. Incorrectly chose the latter, although she knew that it was not a high school.
**R10:** R circled "1" because he had "never heard of any of that other stuff."

## III. Alternative Programs

**Explanation box: The next few questions have to do with alternative programs. An alternative program is one which focuses on helping students stay in and complete school. An alternative program can be a part of a regular high school or it can exist independently, but all offer special services that may not be given in a regular high school program.**

**R1:** R initially skipped the instructions but was unable to answer Q26, she then read the instructions. R did not seem to know the words "alternative' or "independently".
**R2:** R did not understand what an alternative program is.
**R3:** R said that the explanation does not register when read aloud. R was still unable to explain what the intent of the explanation was after reading it silently.
**R4:** No finding.
**R5:** R was unclear as to whether or not she should read the explanation.
**R6:** R had to reread the statement several times and asked for clarification as to the meaning of alternative programs. She recommends deleting the second sentence because its unnecessary and hard to read.
**R7:** R read box, but couldn't immediately paraphrase meaning and had to re-read. Interpretation = "better program"
**R8:** Did not understand term exist.
**R9:** No finding.
**R10:** R did not read box.

### Q. 26 - "Have you ever participated in an alternative program?"

**R1:** See Explanation box.
**R2:** R answered that she had been in an alternative program. A study period that a teacher had recommended she take.
**R3:** R responded that she never participated in an alternative program, although she was not sure what an alternative program is.
**R4:** No finding.
**R5:** R did not read the instruction aloud.
**R6:** No problem answering after alternative programs clarified.
**R7:** R circled "no" instead of code. Said for yes/no responses, just have words (yes/no) and not #'s too.
**R8:** Unsure what an alternative program was, despite reading text in box. He changed his answer between yes and no several times. After settling on no, he correctly followed skip.
**R9:** Reported that the preparatory school she participated in is an alternative program. Being unfamiliar with the school, I am unsure whether this is the case. However, as noted above, this resulted in the school being reported twice as it had been reported in questions 13 through 16 as the school she had returned to.
**R10:** R missed skip instruction here.

### Q. 27 - "When did you enter the most recent alternative program?"

**R1:** R entered 3/4 for month. She considered the GED program she is currently enrolled in as an alternative program.
**R2:** R could not remember the month she entered the program.

**R5:** R considered the Urban League an alternative program. She did not read the instruction aloud.
**R8:** No finding.
**R9:** No finding.

Q. 28 - **"When did you leave the most recent alternative program"**

**R2:** R did not read the instruction aloud.
**R5:** R did not read the instruction aloud.
**R9:** No finding.

Q. 29 - **"Please write below who referred you to this alternative program."**

**R2:** R did not read the instruction aloud.
**R5:** R did not read the instruction aloud. Although at the beginning of the question there is an instruction to "write below".
**R9:** Seemed to have defined referred as including a recommendation by a friend, although also (correctly) said self.

Q. 30 - **"Why did you enter this alternative program?"**

**R1:** R's response "to get GED" does not really answer the question.
**R2:** R did not read the instruction aloud.
**R5:** R did not read the instruction aloud.
**R9:** No finding.

Q. 31 - **"Why did you enter this program?"**

**R1:** R initially missed the instruction to circle one choice on each line.
**R2:** R read through the list of choices but did not consider the "Was not offered" category. She also had a difficult time with the word "peer".
**R5:** R understood rewards to mean about the same as awards.
**R9:** Initially R was unsure what question meant, perhaps because she inserted a why at beginning when reading question. Answered a, b and c by thinking of the prep school she recently attended, but then answered d through h thinking of the high school she used to attend, probably because that was the school that had these programs. When choosing the phrase was not offered she was interpreting this as meaning that the school did have such a program, but she did not ask to participate (or maybe that the program was not directly offered to her). COMMENT: PHRASE WAS NOT OFFERED IS AMBIGUOUS AS IT CAN MEAN EITHER THAT THE SCHOOL DID NOT HAVE A PROGRAM OF THIS TYPE, OR THAT IT DID, BUT NO-ONE ACTUALLY ASKED IF RESPONDENT WANTED TO PARTICIPATE. WORDING SHOULD BE CLARIFIED, E.G., IF LATTER IS INTENDED, WAS NOT OFFERED TO ME WOULD BE BETTER.

Q. 32 - **"Did you participate in any other alternative programs prior to the most recent one?"**

**R1:** R considered Job Corps as an alternative program.
**R2:** No finding.

**R5:** R did not read the instruction aloud.
**R9:** No finding.

## IV. Plans for the Future

**Q. 40 - "As things stand now, how far in school do you think you will get?"**

**R1:** R seemed confused, she chose item #3. In Q42, R answered police officer. WHERE DOES PLOICE ACADEMY FIT IN Q40? ALSO, WHERE WOULD GED GO?
**R2:** Q. 40: R did not understand the question. After we discussed the question R decided on GED and circled "High school graduation only".
**R3:** R did not understand the question and made a guess for a choice.
**R4:** No finding.
**R5:** R did not read through the list of choices. She read through three of the items and picked the third. R did not initially distinguish between "less than years" and "two years or more" when choosing #3.
**R6:** R objected to this question for several reasons. First, she dislikes these type of questions because at 19, how do you know what you're going to do about school once you've dropped out and have kids to raise. She feels researchers have low expectations for dropouts anyway; she said otherwise they wouldn't ask this type of question. She also pointed out that many dropouts may have both realistic and high goals for returning to school---realistic being what person feels is doable and high being what person would like to accomplish. She also recommended that the stem of the question, "as things stand now" be changed to "as things are now" or "right now".
**R7:** R's answer is primarily based on sister's experience in business school. My impresssion is that it was hard for this R to soundly project into the future, and when she does, she bases it on the here/now experience or the experience of others. Asked what she would take in school and she is interested in computers.
**R8:** Read alternatives and said that none applied.
**R9:** R seemed to experience a little trouble giving an answer (finish college), but given the uncertainty she showed on other questions regarding educational plans this seemed speculative, and quite optimistic. COMMENT: GETTING A GED OR HIGH SCHOOL EQUIVALENCY IS NOT OFFERED AS A RESPONSE OPTION--FOR THE DROP-OUT QUESTIONNAIRE AT LEAST THIS SHOULD BE ADDED.
**R10:** R said he circled "6" because he planned to go into military after GED and would go to college on GI bill. PERHAPS MILITARY SHOULD BE ADDED AS RESPONSE CATEGORY?

**Q. 42 - "Write in below the name of the job or application that you expect or plan to have when you are 30 years old. Even if you are not sure, write in your best guess."**

**A. Which of the categories below comes closest to describing that job?**

**R1:** R chose #7. From answers to other questions, #12 would have been the right answer; apparently she did not read through the list. IS THE LIST TOO LONG?
**R2:** R read the occupational categories in caps aloud to "FARMER" then went back and circled #1 for "CLERICAL".
**R3:** No finding.
**R4:** No finding.
**R5:** R did not read the instruction aloud. She read most of the choices aloud and

others silently. R selected "other" to categorize mortician.
**R6:** R's reaction to this question was that it was too hard to predict that far into the future, too many categories and too hard to find type of job within the categories. She classified her expected job as "other" because she couldn't find a category that fit becoming an AIDS counselor. Other comments-- she asked why its the R's responsibility to find and circle the appropriate category to describe the job, and not someone in the research office, since R is recording the type of job, felt this effort is very time consuming, especially for poor readers; the "when you are 30 years old" part of the question gets missed as she had to read it twice to make sure she understood that's what it said, would recommend underlining or bolding the information; she asked why there are two Professional categories and what is the difference between them; she asked why there's a line to specify at "other" when R has already recorded the job above.
**R7:** Made choice because of current interest in computers. R said she was sure about answer. Computer operator job functions include repairing and programming computers. When asked how found job in list- looked at operative first, even though "operative" is in the middle of list. R said she understood words on list, but when probed proprietor, R said part-owner. When probed her choice of computer operator with other categories such as clerical, manager/administrator, she said that she didn't feel what she would be doing would fit into those categories, even though they might be applicable for what others would do as computer operators. When I probed why not technical, she acknowledged that computer programmer was part of the list, but that she would do more than programming. It is clear that although R understands the colloquial definition of these job headings, their specificity and use here may not integrate into R's thinking. Either that, or the examples are misleading if R's specific job doesn't match exactly the examples shown. This R thinks of a computer as a machine, not all that different from a print press let's say, and not a machine that requires technical training and knowledge.
**R8:** Chose appropriate category, although may well not have done if the examples for craftsman did not include specific occupation he wanted to do (carpenter). Wrote yes by word, rather than circling number. He continued reading categories below this.
**R9:** R wrote in two related jobs with very different training requirements (vet and vet assistant). Q. 42a - R looked for a mention of the specific job and not finding it wrote in under other. When asked to look again at the categories selected professional, on the grounds that vets are similar to physicians; thus she chose the appropriate category for the more prestigious of the two occupations reported on Q42 (presumably vet assistant is a technical job).
**R10:** R had much trouble with this Q. Was unable to think about 10 yrs in the future. This question brought out a lot of fatalism: R kept saying he'd be dead by then. Arrived at answer only with much probing and prodding. Think he would have left it blank had I not been there...

## IV. "Language Use" (administered only to the respondent for whom English was not the native language)

**Q. 46. - "What is your native language (the first language you learned to speak when you were a child)?**

**R6:** R reflected on this question because (1) in her mind most researchers think that all Hispanics speak Spanish as their native and first spoken language, so she worried that her data would appear suspect if she answered English, and (2) she actually spoke Spenglish because both languages were spoken simultaneously in her household. She recommended rewording the question to read "when you were a child, what was the first language you learned to speak (your native language)? She feels this way people focus better on what they learned to speak as opposed to what they speak now.

## V. "Money and Work"

Explanation box at top of page 26: **We are interested in learning about the kinds of jobs you have held, the hours you worked and your income from these jobs, the level of your job satisfaction and the relation of your training and education to your work experience. This information will help us better understand the movement of young people into the world of work and the reasons for changes in job situations.**

**R1:** Read aloud, the explanation seemed a bit burdensome.
**R2:** No finding.
**R3:** The explanation was not clear to R. She thought it was about learning and jobs.
**R4:** No finding.
**R5:** No finding.
**R6:** R felt language too elitist--"who do they think they're talking to with words like "job satisfaction" and "relation to your training and education" when most don't have no training and you wouldn't be interviewing them if they had an education."
**R7:** R acknowledged reading paragraph. Said didn't really help understand what the questions in the sections were about. R probed for clarity as to why not. At this point in interview, R was distracted. Focused interview by asking her to point out key words/phrases in explanation that helped/would help in her understanding: Income, job satisfaction, mention of jobs. Related paragraph statements immediately to job experiences. Began discussing specific job and what events happened there. So, the statement did key the R's focus on thinking about her work experiences and income.
**R8:** Read text but appeared not to understand it.
**R9:** R read and understood this.
**R10:** Read it.

**Q. 53 - "Whether or not you already have a job, were you looking for a job last week?"**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** R did not read skip instructions aloud.
**R5:** R did not read the instructions aloud and missed the skip to Q55.
**R6:** Thought reversal of clauses would make the question more clear.
**R7:** Probed for R's definition of last week, still, Sunday through Saturday. Asked whether R saw/read stem of questions - "yes". R also noted skip pattern.
**R8:** Correctly answered.
**R9:** Had to read question twice before understanding it, I think because she was confused by phrase: whether or not ... Reverted to circling word (No) rather than number, and perhaps because of this missed the skip instruction.

**R10:** No finding.

Q. 54 - "Have you done any of the following in the last week to find a job?"

**R1:** When R was probed about private employment agency she thought it meant finding work in a private home. R chose "Other" to classify just calling around. There was also some question about the difference between item f and g. She missed the skip to Q56.
**R3:** R did not read instruction but did read through the list. She only circled #2 - Yes for those choices that applied. Of the three choices she selected, she initially picked f. then changed to g. The distinction she made between the two is that to answer ads in a newspaper means to send a resume, to look in the newspaper is to look and not respond.
**R4:** R chose g. over f. because she only looked at ads and did not respond to any.
**R5:** When asked by the interviewer, R thought that a private employment agency is a privately owned business. R missed the skip to Q60.
**R8:** Correctly followed instructions, and selected category e correctly. However, he was unable to find a phrase that indicated two other job search activities he had done last week (applied at a shop for work and applied for a city summer job). He initially circled yes for Other, then changed his answer to no.
**R9:** Should have skipped question.
**R10:** OK; circled YES for other; probing elicited the fact that he had checked with a temporary service and did not consider that "directly with employer." Missed skip.

Q. 55 - "Why weren't you looking for a job last week?"

**R2:** R did not read the instruction aloud.
**R5:** R did not read the instruction and answered yes/no to each choice.
**R6:** She thought Qs.55 and 53 should be combined somehow so that if R is working, doesn't have to answer why he/she not looking. She thinks "that's stupid".
**R7:** Recorded answers in "other" - she didn't note CIRCLE ONE instruction nor recognize that her answers fit appropriately in two of the answer categories (#2,#3). R had difficulty choosing one main answer when probed, but finally decided on #3. Attempted to probe R's future decision making ability by asking how she intends to resolve child care problems to attend school in May for GED. Although she has some ideas, plans are very contingent on will of others, have not been really thought about or investigated by R and point to her dependence on others to help make future decisions. Attempted to test R's understanding of answer categories by using different vignette (main reason to go part-time) - R was able to distinguish that "other" best suited that main reason.
**R8:** Missed skip, read through items and wrote no on the other line.
**R9:** Commented that question seemed repetitious, looked at previous page and noticed skip on Q54 (rather than the skip she should have followed, on Q53) and (incorrectly) did not answer this question.

**Q. 56 - "How many jobs have you held since you last left school?"**

**R1:** R was confused about the skip to Q.57.
**R2:** R was not sure if she should count her current job.
**R3:** Initially R did not count her current job.
**R4:** No finding.
**R5:** R did not read the instructions aloud and missed the skip to Q60.
**R6:** R recalled first job held first and then went forward---not current job and then backward. I asked her why she used this approach and she said that she got her first job the same month her son was born (significant event recall). She actually counted the jobs on her hands to keep track and counted them twice to make sure she didn't forget any. I asked her what school she used to frame her answers and she said high school and not the trade school. She also said some Rs may not know how to count their "hustles"--doing hair, selling merchandise on the street, babysitting, etc.--not illegal acts just ways to make money that aren't formal jobs.
**R7:** R has had only one regular job since leaving school, but has other income source, hairdressing, that she does occasionally. Considers this more a hobby than a job. If intent of question or section is to arrive at income sources, will need additional question to probe sources like hairdressing, babysitting, car repairs. R was able to distinguish time frame - since leaving school.
**R8:** Initially answered none, probably because he read it as jobs held now. Then changed answer to one (probably correct).
**R9:** R answered by thinking in terms of when she left high school. COMMENT: STRICTLY SHE SHOULD HAVE THOUGHT OF THE PREP SCHOOL SHE RECENTLY LEFT, ALTHOUGH I THINK INTENT OF THE QUESTION WAS BETTER SERVED BY ANSWERING FOR THE HIGH SCHOOL. She said she had done about 20 jobs, but was including some quite casual work, such as babysitting.

**Explanation box at top of page 28: NEXT WE WOULD LIKE INFORMATION ABOUT TWO OF THE JOBS YOU HAVE HELD SINCE YOU LEFT SCHOOL. WE WOULD ALSO LIKE TO KNOW ABOUT ANY PERIODS OF TIME YOU WERE LOOKING FOR WORK BETWEEN JOBS. IF YOU HAVE BEEN IN THE MILITARY, PLEASE CONSIDER YOUR ENTIRE MILITARY EXPERIENCE AS ONE JOB.**

**R1:** R was unsure whether to read this explanation. Possibly because it is on a separate page that is mostly blank.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R5:** No finding.
**R6:** R initially missed reading this because of the way it was xeroxed (probably should have been on the same page with Q.57 since previous questions tell R to skip to Q.57.) Her understanding of the statement was good.
**R7:** R acknowledged reading box, even though skip at Q. 56 says go to Q. 57, and instruction is in between. However, not really sure she actually read it, based on her ability to paraphrase and emphasizes jobs. Probed her understanding of military experience, she grasped that separations from and return to military would count as one job, but said that other R's probably wouldn't know.
**R9:** R read and reported that it was clear.
**R10:** Skip at Q56 says "Go to Q 57," so that's where he went, skippping this box entirely. This instruction should be part of Q57.

**Q. 57 Instructions: "CURRENT OR MOST RECENT JOB HELD SINCE YOU LEFT SCHOOL. IF YOU HAD TWO JOBS AT THE SAME TIME, ANSWER FOR THE JOB YOU HAD THE LONGEST."**

**R1:** No finding.
**R2:** R was not sure how to answer the question. She is currently working as a secretary for the Urban League. The answer she gave is public business.
**R3:** No finding.
**R4:** No finding.
**R6:** R was a bit confused about the difference between current and most recent job. Once explained, she felt that other Rs would have same problem. Recommended that statement should read something like: "Current job where you are working now or the last job you worked at since you left school".
**R7:** Probed for R's understanding of current/most recent job - she understood and gave examples. Also probed for understanding of two jobs at the same time. She understood. However, R thinks "current" should be referred to as job person has now.
**R8:** Read them, unclear if understood.
**R9:** Confusion described for Q57A probably reflects uncertainty about the meaning of these instructions.
**R10:** Was dumbfounded by the language here. Wanted to focus on second sentence even tho it didn't apply to him.

**Q. 57A - "What kind of job or occupation did or do you have? (for example, salesperson, waitress, secretary, assembler, etc."**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** R wasn't sure if she should describe what a "peer educator" is here or just record the type--seems the natural inclination is to do both. The "did you or do you have" alternate word choices and others like them were troublesome to her. Also, this is where R was adamant about the use of degrading examples--"don't they think we can do work besides salesperson or waitress, or work in places besides Pay Less or McDonald's"--she strongly advocates getting rid of the examples or at least including other types.

**R7:** Probed whether "cashier" title also included sales/waiting on customers or something else. At this point said no, just ringing up orders. However, later at Q. 57c, did specify other tasks done, but title she said was cashier.
**R8:** Wrote port, meaning porter.
**R9:** Although R correctly reported her current job, the did you or do you have left her uncertain whether question was asking about her current or last job.
**R10:** R had a lot of trouble titling his job, since he did a little bit of everything. One major pitfall: R worked for temp service, so he did different things on different days. He probably should have been instructed to write in "temp" or some such.

**Q. 57B - "What kind of business or industry was this job in (for example, retail shoe store, restaurant, electronic assembly plant)?"**

**R1:** No finding.

**R2:** R did not know how to answer this question. She is currently working as a secretary for the Urban League. The answer she gave is public business.

**R3:** No finding.

**R4:** No finding.

**R6:** Classifying industry is really impossible for Rs--still not sure I probed enough to get all the info needed. R had a difficult time understanding my probes and the need to find out more about the "health care project" she recorded. I think she did a good job of arriving at an answer without benefit of probes. R recommends putting the question mark after industry or change question to--"What type of company is this or Where are/were you employed"? Again, she thinks the researchers should be smart enough to figure it out if they have a name of the company or the type of job it is and what the duties are. I explained the traditional problems we encounter with this question, even when interviewer administered, and R said "sounds like the question shouldn't be asked".

**R7:** R asked about meaning of question. Asked her to paraphrase meaning, and she essentially said "where do you work/what's its name/ and what kind of place is it?" She then recorded franchise name. I probed, she then classified as a snack shop because sold only candy, snacks and ice cream.

**R8:** Was a porter for a Woolworth restaurant. Wrote restaurant rather than shop, perhaps because this was specified in question.

**R9:** Correctly answered.

**R10:** R could only answer this when I paraphrased as, "What did they make or do?"

**Q. 57C - "What were your main activities or duties on this job (for example, selling shoes, waiting on tables, putting computer boards together)?"**

**R1:** No finding.

**R2:** No finding.

**R3:** No finding.

**R4:** No finding.

**R6:** R had no problem answering question and had recorded her main duties before any probing. I did ask her to explain what "educating" means.

**R7:** Probed for description of cashier duties and this is when other duties than ringing up orders were described. It will be very difficult to get respondents to record more than one or two word descriptions or simply repeat their job titles as R did.

**R8:** Activity specified (washing dishes) seems inconsistent with job specified (porter).

**R9:** Correctly reported.

**R10:** R did not see any difference between this and 57a.

**Q. 57D - "When did you start working at this job?"**

**R1:** R entered 9/14 for month, meaning September 14.

**R2:** No finding.

**R3:** No finding.

**R4:** No finding.

**R6:** R had no recall problems when arriving at start date. She likes this job and remembers well when started.

**R7:** No recall problem. Tied to other significant life event.

**R8:** No finding.

**R9:** R showed some uncertainty about when started job, partly because she started doing it quite gradually. Her answers to later questions about first job (Section V, Q. 57D,E) suggested that current job started 4/90, rather than 4/89

as she reported.
**R10:** No finding.

**Q. 57E - "When did you start looking for this job?"**

**R1:** R entered 91 for year, although she said 90 and that appeared to be what she meant.
**R2:** No finding.
**R3:** No finding.
**R4:** R was not looking for a job, it was offered to her. She entered the same dates as those in Q57d, the dates when she started working the job.
**R6:** Again, no recall problems.
**R7:** Had to reread question because she automatically assumed the question asked for the date she left job. Said that questions were in an order that did not make sense.
**R8:** Seemed to answer in terms of when he went for the job interview, rather than when he started looking for work of this type.
**R9:** R did not answer, saying that question did not apply as she never looked for job (just started working with boyfriend as he started his business).
**R10:** No finding.

**Q. 57F - "Do you still have this job?"**

**R1:** R entered 12/15 for month, meaning December 15.
**R2:** R did not read instructions.
**R3:** No finding.
**R4:** No finding.
**R6:** Still working at job--also note that she circled "yes" instead of code.
**R7:** No problem.
**R8:** Wrote 90 for year, although I think he left in '91.
**R9:** Answered correctly and followed skip, although R started reading Q57h before remembering skip instruction.
**R10:** Missed skip.

**Q. 57G - "Why did you leave this job?"**

**R1:** Did not read item instruction aloud.
**R4:** R left the job because it was seasonal work. She selected "Other" to answer the question rather than #1 which includes temporary job. Her reasoning was that when a temporary job ends it is over forever. A seasonal job ends for the season but a person can get the same job back the next season.
**R7:** Answered without hesitation -- pregnancy was the primary reason, although she also cited not getting along with her boss and low pay.
**R8:** Wrote fired on line by appropriate category.
**R10:** No finding.

Q. 57H - "Were you without a job AND looking for work right after you left this job?"

**R1:** No finding.
**R4:** No finding.
**R7:** R was somewhat confused how to answer (double-barrelled question). Not sure how to respond.
**R8:** As no other job reported, answer was inconsistent with date given when left job.
**R9:** See comment for Q. 57F.
**R10:** Initial trouble with "weeks" since he had been out of work for months at a time.

Q. 57I - "How much do/did you earn per hour when you first started this job?"

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** R was able to find appropriate range easily. I asked her what she would have done if she made less than $3.75/hr when she started job and she said she would have written in amount in margin. I asked her if she thought the bottom range was too low, too high, or about right and she said "about right, because most folks earn at least that much, even in the Latino community."
**R7:** No finding.
**R8:** Wrote hourly rate on line by the appropriate category.
**R9:** R experienced much difficulty answering as she works for her boyfriend's business and does not get paid. Instead they have a joint bank account and she takes out money when it's needed. Eventually chose category 02, because their employee earns that much, so that is what she thought she would get if she was receiving hourly wages.
**R10:** No finding.

Q. 57J - "How much do you earn currently, or did you earn just before you left this job?"

**R1:** Had difficulty reading this item.
**R2:** R's initial response to the question was "does currently mean now?"
**R3:** No finding.
**R4:** R initially selected #1 thinking that it included her wage of $4.75. She realized her mistake and changed her choice to the correct choice #2.
**R6:** Again, R had no trouble finding appropriate range.
**R7:** No finding.
**R8:** Same as 57I.
**R9:** Selected category 02, using same reasoning as for Q57I.
**R10:** No problem.

Q. 57K - "About how many hours a week did or do you usually work in this job?"

**R1:** No finding.
**R2:** No finding.

**R3:** No finding.
**R4:** No finding.
**R6:** R had no trouble arriving at number of hours. She did comment again that she was bothered by the "did or do you". Recommends "do/did" showing present tense first.
**R7:** First recorded 20 hours per week because there were a few weeks she worked 4 hours instead of the usual five. THE "usually" NEEDS TO BE MORE DISTINCTIVE.
**R8:** In space for the tens wrote in number of days he worked per week (5) and in space for single numbers wrote in hours. Asking for time he started and finished work indicated that he had worked 5.5 hours per day, rather than the 4.5 that he reported.
**R9:** First reported 41 hours, then 47+ hours. Was answering by thinking of how long shop was open each week. When asked directly, R reported working 16-18 hours overtime in a typical week; none of the overtime hours were included in the 47 hours reported.
**R10:** No finding.

**Q. 57L - "How did you find this job?"**

**R1:** R did not read through all the choices. Chose "other" and wrote "walked in". Probably should have circled #5. Did not read instructions aloud.
**R2:** R did not read the instructions.
**R3:** R did not read the list aloud but skimmed through and chose relative.
**R4:** No finding.
**R6:** Initially, R wanted to circle two answers--codes 6 and 7. Circumstances were that a friend of her aunt's told aunt and aunt in turn told R. When I probed who specifically told R, she said aunt and then circled code 6. Brings up issue that some Rs may circle more than one answer because they feel they had multiple sources.
**R7:** Had a hard time making choice. Said that other Rs might have trouble in noting and making the selection -- several referrals to the same job.
**R8:** No finding.
**R9:** Despite instructions, selected two categories, relative and friend. However both categories referred to the same person, her boyfriend.
**R10:** No problem.

**Q. 57M - "In this job are/were you..."**

**R1:** R worked at a drug store and a fast-food restaurant. She did not think that those were private companies. R felt that none of the categories covered her jobs. Later on, she could not suggest a better phrase. Reordering the categories might help -- own business, family's business, someone else's business or corporation, government.
**R2:** R was not sure of the category to choose for her job. She decided on "Private Company".
**R3:** R had a problem with this question. After much deliberation she chose "Private Company".
**R4:** No finding.
**R6:** For current job, R was confused as to her status because the program she works for is funded partially by the city and partially by a private hospital.

When I probed who pays her, what type of check she receives, or if she knows which (city/hospital) provides the funding for the staff's salaries, she thought a minute and said she was pretty sure the city did. If I hadn't been there, I asked her what she would have circled and she said either both codes 1 and 2 or she would have left it blank and written something in the margin. R feels that most Rs wouldn't know how to classify their jobs. She also thought that categories 4 and 5 should be before category 3--how likely is it that a dropout at 18 or 19 will own his/her own business?
**R7:** This question was a real problem for R. R didn't ask for clarification when answering but in talking about the question, she admitted not understanding the differences between private companies, government, family business with/without pay. R didn't understand family business even after several attempts to clarify. Confusion focused on the ambiguity of "own family owns" or "anyone's family owns".
**R8:** Read through categories, did not find one that he thought applied and gave no answer.
**R9:** Said "self-employed". Strictly wrong, as her boyfriend is the owner. R selected this category as she considers that she plays an equal role in all aspects of the business, including the making of decisions.
**R10:** No finding.

Q. 57N - "How did you learn to do this job?"

**R1:** I asked if an employer-sponsored training program is the same as on the job training? Again R did not read all the options.
**R2:** R read through the first three choices and then fit her job into one of the three.
**R3:** No finding.
**R4:** R was unsure of what the word apprenticeship means.
**R6:** Initially, R didn't see instruction to circle all that apply--shift in instructions from previous questions. She asked me how she was supposed to code answer if more than one and I probed by asking her what instruction said. She felt that instruction should be larger and bolder. She felt most Rs would be able to understand categories, would read entire list, and code all that apply, if they see instruction.
**R7:** R had been a cashier at friend's family restaurant where she learned to work the cash register. Learned other parts of the job from co-worker but primarily self-taught.
**R8:** He selected correct category, but wrote yes rather than circling number.
**R9:** Incorrectly selected category c (apprenticeship), although she got no formal training and learned by her boyfriend/company owner showing her what to do.
**R10:** No finding.

Q. 57O - "Have you held other jobs since you left school?"

**R1:** Did not read skip instructions aloud.
**R3:** R initially missed the word "other" when she read the question the first time.
**R4:** No finding.
**R6:** Again, R circled Yes instead of code.
**R7:** Immediate response "no, because of no one to take care of my son".
**R8:** Incorrectly answered yes, reporting a job he had had while still at school.

**R9:** No finding.
**R10:** R misinterpreted this. Need phrase "Besides the job you just told us about..." or some such.

**Q. 58 Instructions - ANSWER THIS QUESTION FOR THE FIRST JOB YOU HAD AFTER LEAVING SCHOOL. IF YOU HAD TWO JOBS AT THE SAME TIME, ANSWER FOR THE JOB YOU HAD THE LONGEST.**

**R1:** No finding.
**R2:** R had to re-read the instructions to answer question A.
**R3:** The instructions did not make much sense to R until after she re-read it a couple of times. It was not clear if she actually knew what the question was after re-reading it.
**R4:** R confused Q57 with Q58 so the answers for questions Q57A through Q57N are for the most recent job she has had since leaving school. The answers for Q58A through Q58N are for the same job as that for Q57A through Q57N.
**R6:** R asked whether first job could include a job held when dropped out. If so, R recommends that it be part of instruction and that "first job" be underlined for clarity.
**R8:** No finding.
**R9:** Misunderstood question, and answered for job she held the longest, even though it was not the first job she held after leaving school. This longest job started three years and many jobs later. Questions 58A through N were then answered by reporting for this incorrect job.
**R10:** Same as comment for instruction box above.

**Q. 58A - "What kind of job or occupation did you have (for example, salesperson, waitress, secretary, assembler, etc.)?"**

**R1:** No finding.
**R2:** R was not sure if she gave the correct answer to the question. It was not clear to her which the question was asking about.
**R3:** No finding.
**R4:** No finding.
**R6:** R's recommended changes-- same as Q57.
**R9:** No finding.
**R10:** Same as above.

**Q. 58B - "What kind of business or industry was this job in (for example, retail shoe store, restaurant, electronic assembly plant)?"**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** Same as above.
**R9:** No finding.
**R10:** Same as above.

**Q. 58C - "What were your main activities or duties on this job (for example, selling shoes, waiting on tables, putting computer boards together)?"**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** Same as above, when I probed further about duties, R expanded description -- she felt most R's would limit descriptions unless prompted.
**R9:** Described job well.
**R10:** Same as above.

**Q. 58D - "When did you start working at this job?"**

**R1:** R could not remember the exact date.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** No recall problems. Tied to significant life event.
**R9:** Answered 6/89, but was uncertain whether it was June or July.
**R10:** No finding.

**Q. 58E - "When did you start looking for this job?"**

**R1:** Initially R could not remember the date. She did recall the month but gave the wrong year.
**R2:** No finding.
**R3:** R's response was the same as Q57E, that she did not look for the job it was offered to her. It may be that she confused the two in answering the questions and did not want to admit it, or she actually was offered both jobs.
**R4:** Same as Q57E, job was offered to R.
**R6:** Same as above.
**R9:** Gave same month as the previous question on when started job. R said she was not looking for a job of this type any earlier because she was working at another job. However, her comments suggested that she was looking while working at previous job.
**R10:** No finding.

**Q. 58F - "When did you leave this job?"**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** Same as above.
**R9:** Did not answer, did not ascertain why.
**R10:** No finding.

**Q. 58G - "Why did you leave this job?"**

**R1:** No finding.
**R2:** R did not read the instructions or the entire list. She picked the first item "job ended" because she was fired.
**R3:** R's answer was that job was not interesting written in the "other" category. Possibly #3 would have been the more appropriate choice.
**R4:** Same as Q57E, seasonal job.
**R6:** Very clear in R's mind why--no babysitter. She was somewhat offended that there wasn't a category for that, especially as that's the reason most female dropouts have for not working, returning to work, or not staying employed.
**R9:** Wrote in a reason that comments suggested was correct. COMMENT: WRITTEN-IN ANSWER IS A COMMON REASON FOR QUITTING A JOB (NOT GETTING ALONG WITH EMPLOYER)--MIGHT MAKE IT ONE OF RESPONSE OPTIONS.
**R10:** R did not realize that response categories went with Q. He thought it was open ended and was going to write his response in the white space at the end of the stem.

**Q. 58H - "Were you without a job AND looking for work right after you left this job?"**

**R1:** R went to Job Corps. How is Job Corps classified, as training or as a job? She also missed the number of weeks looking for a new job.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** Definitely a double-barreled question--R can and did have two responses. "Yes, was without a job" and "No, not looking for work right after leaving this job". Decided the answer was no because she really wasn't looking and to say "yes" would require entering # of weeks.
**R9:** Answered no, started answering the follow-up question to a yes response, but noticed error.
**R10:** Months?

**Q. 58I - "How much did you earn per hour when you first started this job?"**

**R1:** No finding.
**R2:** It was not clear to R if this question was asking about the job she has now or the one from which she was fired.
**R3:** No finding.
**R4:** No finding.
**R6:** Initially, R didn't know which job question was referring to--the current job or the one at the grocery store (first job). Then she remembered that she had already answered these questions about her current job. I asked why she was confused and she said that having Q.58H in between sort of threw her off because that question asked about looking for work. She also said the word "that" should replace "this" when referring to job because it made her think about her current job. R had no problem finding appropriate range once the job was clarified.
**R9:** No finding. COMMENT: ORDER OF QUESTIONS DOES NOT FOLLOW CHRONOLOGY, PERHAPS ADDING TO COMPLEXITY AND RESPONDENT CONFUSION. Q58I AND J, ON PAY, COMES AFTER QUESTIONS ABOUT LEAVING THE JOB; THEY SHOULD AT LEAST COME BEFORE 58H.
**R10:** No finding.

Q. 58J - "How much did you earn just before you left this job?"

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** R knew which job for this question and had no problem finding appropriate range.
**R9:** No finding.
**R10:** No finding.

Q. 58K - "About how many hours a week did you usually work in this job?"

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R6:** R's recall on this involved how many hours per day multiplied days she worked per week. Had to point out the word "usually" in the question when she said that there were some weeks she only worked four days because of babysitting problems. She thought usually should be more visible because other Rs may have that same problem given that many work part-time hours.
**R9:** Hours reported was just one hour different from the hours arrived at by calculating them from the starting and leaving times reported.
**R10:** No finding.

Q. 58L - "How did you find this job?"

**R1:** Did not read item instruction aloud.
**R2:** R read through the list to #4 then jumped to #7 and circled it.
**R3:** No finding.
**R4:** No finding.
**R6:** R had no problems with this question.
**R9:** No finding. COMMENT: AS WITH Q58I QUESTION L AND M AND N PROBABLY SHOULD BE ASKED BEFORE QUESTIONS ON LEAVING JOB.
**R10:** OK.

Q. 58M - "In this job were you..."

**R1:** R's response same as Q57M, that none of the categories were appropriate. Again, did not think big corporation counted as private company.
**R2:** R had a serious problem with this item. She worked at Safeway but picked #2 because of "local" in the choice and Safeway store is a local company.
**R3:** No finding.
**R4:** No finding.
**R6:** Again, some confusion on how to code. The store was owned by a family, so she interpreted that to mean code 4 instead of code 1. R feels the categories should be more specific, use examples, etc. Definitions of what a private company, government employee, etc. are necessary. She really recommends that the question not be asked.

**R9:** R somewhat unsure, but after some thought answered correctly (private company). COMMENT: QUESTION HAS SEVERAL POSSIBLE AMBIGUITIES: WHEN WORKING FOR A SINGLE INDIVIDUAL, LIKE A PHYSICIAN, CALLING IT A PRIVATE COMPANY SOUNDS A LITTLE TOO GRAND; FAMILY BUSINESS OR FARM PRESUMABLY REFERS TO RESPONDENT'S OWN BUSINESS/FARM, BUT RESPONDENT MAY CHOOSE THIS CATEGORY IF HE/SHE WORKS FOR ANY FAMILY BUSINESS. OWN FAMILY BUSINESS MAY BE CLEARER.
**R10:** No finding.

**Q. 58N - "How did you learn to do this job?"**

**R1:** No finding.
**R2:** R was not sure of what an apprenticeship is.
**R3:** R did not read through the list. She recognized the choice from the list in Q57N.
**R4:** No finding.
**R6:** This time R saw instruction to circle all that apply. No problem with question.
**R9:** As with Q 57N, R incorrectly reporting learning on-the-job as an apprenticeship.
**R10:** No finding.

**Q. 59 - "How much of the money you make is spent on each of the categories listed below?"**

**R1:** No finding.
**R2:** No finding.
**R3:** R missed the instruction to circle one on each line and only circled #1 or #2 for the choices that applied. "None of it" was not selected for any of the other choices.
**R4:** No finding.
**R6:** Initially, R didn't answer category e. When asked why, she said didn't apply to her so she left it blank. After answering category a, she asked whether "to buy things" included purchasing clothing, and I said yes. I asked what else she spends money on and she added babysitters and getting her hair done.
**R7:** Asked for clarification of what was included in several categories. No clue as to how to code money spent to care for her child. Said she also spends money on babysitter and video rentals. Also thought last category was stupid for dropouts -- makes them feel guilty if they're not.
**R8:** For three of the four items that he spent no money on he left blank, rather than circling a number. Chose some of it for a, although he probably should have selected most of it.
**R9:** R incorrectly answered for longest job done, rather than current job. COMMENT: THIS IS NOT SURPRISING, GIVEN PREVIOUS QUESTIONS CONCERNED FIRST/LONGEST JOB, AND Q59 DOES NOT MAKE EXPLICIT THAT CURRENT/MOST RECENT JOB IS NOW BEING REFERRED TO. Rather inconsistently, R reported that most of the money is spent on both rent and food (c and d).
**R10:** No finding.

**Q. 60 - "Have you participated in and completed any of the following types of training programs or courses?"**

**R1:** R had trouble with the phrases "formal apprenticeship" and "correspondence course". She circled No not #1 for item a. but did circle #1 for govt. training programs and correspondence courses. For item d. R circled "If yes" not #2 and did not circle anything for "If yes ... completed" but did enter a date for "Date completed". The program she wrote in for item d. was Job Corps. She also mentioned her GED program.
**R2:** R had a difficult time with this question. She did not know what "Formal Registered Apprenticeship" or "Correspondence courses" were, and could not think of any "Government training programs" when asked. She did complete a training program at the Urban League which she entered in "Other" and completed the correct sequence of options. She only circled #1 for "Correspondence courses" and nothing for the other choices.
**R3:** R was very confused by this question. She was unsure of what participated meant and missed the instructions; "Have you participated?". She also did not know what "Formal Apprenticeship" and "Correspondence courses" referred to. She did give an affirmative answer to "Government training program" and circled #1-No not completed. The program she was in was the Urban League.
**R4:** R knew how to complete the grid, but was not sure what "correspondence course" meant. She considered the training program at the Urban League as a "Government training program" and completed the item sequence correctly.
**R5:** R considered the Urban League as a government training program, and with some direction successfully completed the required sequence of items. Upon reading the choice of "Correspondence course" she asked the interviewer what it was. There was also some question as to R's familiarity with the concept of "Formal Registered Apprenticeship".
**R6:** Most difficult question for R during interview. The format was confusing to her and the categories were very unclear. She didn't know what items a and c are, so she automatically answered "no". She circled the "no" instead of the code on the line for item a. For item b, she asked what type of government training because her current job involved some training by the city. This question really does cry out for definitions and examples.
**R7:** R totally confused by question format and specific types of training programs or courses being referred to in the categories.
**R8:** Probably answered correctly. Did not understand term correspondence course.
**R9:** Did not understand what a referred to, but correctly answered no. Said yes to d, thinking of on-the-job informal training, and as she was thinking of many jobs was unable to answer the have you completed and date completed follow-up questions.
**R10:** R had no idea what "a" or "c" meant and took a lot of time trying to figure out format before even attempting to answer first column. Needed a lot of help with this format. Also, did not pick up that question was focusing on "TRAINING" programs.

## VI. Drug/Alcohol/Family/Religion

**Explanation box at top of page 46: QUESTIONS 84 THROUGH 87 ARE VOLUNTARY. WE HOPE YOU WILL ANSWER EVERY QUESTION, BUT YOU MAY SKIP ANY QUESTION YOU DO NOT WISH TO ANSWER.**

**R1:** No finding.

**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R5:** No finding.
**R6:** Meaning clear to R.
**R7:** R understood significance of word "voluntary". Said she thought it was good to offer R option of answering, may get more truthful answers.
**R8:** Read it. Unclear if understood.
**R9:** Read and appeared to understand.
**R10:** Read it.

**Q. 84 - "How many cigarettes do you usually smoke in a day?"**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R5:** No finding.
**R6:** R had no problem with the question nor did she hesitate to answer it.
**R7:** Has never smoked and thought there should be a category for that.
**R8:** No finding.
**R9:** No finding.
**R10:** No finding.

Definition box before Q. 85: **IN THE QUESTIONS THAT FOLLOW, ALCOHOLIC BEVERAGES INCLUDES BEER, WINE, WINE COOLERS, AND LIQUOR.**

**R1:** No finding.
**R2:** No finding.
**R3:** No finding.
**R4:** No finding.
**R5:** No finding.
**R6:** R had no problem understanding.
**R7:** R understood purpose of box and language in box. Thought other Rs would do better if added "including mixed drinks" to liquor description.
**R8:** No finding.
**R9:** No finding.
**R10:** No finding.

**Q. 85 - "On how many occasions (if any) have you had alcoholic beverages to drink?"**

**R1:** R initially did not read instructions. If the R enters 0 occasions for use in lifetime should there be a skip out of use for the last 12 months and 30 days? Maybe reversing the order is possible.
**R2:** R read the question but initially missed a through c.
**R3:** R read the question and attempted to answer it as an open-ended question. She then went back and answered a. through c. appropriately.
**R4:** R did not answer for use in "...12 months" and "...30 days" although she reported some use in her lifetime.
**R5:** R responded that she used alcohol on 0 occasions in her lifetime and then

skipped the items asking about use in the las 12 months and in the last 30 days.
**R6:** R had no hesitation answering nor difficulty with format. I probed about time frame for lifetime ("since I been here"), last 12 months ("from April to April"), and the last 30 days ("since the first of this month"). She also understood that occasions means separate times not drinks. I asked whether other Rs would answer truthfully and she said yes unless they're heavy drinkers.
**R7:** No finding.
**R8:** For some reason R only answered part c. probing indicated that he recalled six occasions in last 12 months. He wanted to report one occasion for c, but circled the 0 occasion category.
**R9:** a answered correctly (20+) but b seemed an underreport (3-19 times, yet she reportedly drank alcohol at least two nights each week). Further evidence that b was an underreport is provided by the same frequency category being chosen for the much shorter reference period in c.
**R10:** Ok with categories, but would have omitted answers to b and c.

**Q. 86 - "Think back over the LAST TWO WEEKS. How many times have you had five or more drinks in a row?"**

**R1:** No finding.
**R2:** R reread the instruction and was able to answer the question.
**R3:** No finding.
**R4:** R did not answer for use in "...12 months" and "...30 days" although she reported some use in her lifetime.
**R5:** No finding.
**R6:** No problem with this question. R said people who haven't had a drink in the last month shouldn't have to answer this question because it implies that they were lying at Q.85. Probed about two-week time frame.
**R7:** No recall problem. Two distinct memories.
**R8:** Wrote yes on line by appropriate category.
**R9:** Missed the five drinks in a row phrase, and reported number of occasions she had drank any alcohol in last two weeks.
**R10:** R thought "shot glass of liquor" was funny.

**Q. 87A - "On how many occasions (if any) have you used marijuana (pot) or hashish (hash, hash oil)?"**

**R1:** See question 85.
**R2:** No finding.
**R3:** R did not finish reading the question, but rather vigorously said no to any marijuana use.
**R4:** R did not answer for use in "...12 months" and "...30 days" although she reported some use in her lifetime.
**R5:** R responded that she used marijuana on 0 occasions in her lifetime and then skipped the items asking about use in the last 12 months and in the last 30 days.
**R6:** No problem nor hesitation with this question. Pot makes her "silly" and she doesn't use any more. She offered "that's how I got pregnant the last time fooling with it". I asked whether she thought other Rs would answer--"they'll answer but they may not tell the truth because it looks bad for dropouts, that's what people think of us, that we use drugs."
**R7:** Did use on occasion when she cut school with friends and to ease morning sickness, but never a heavy user.

**R8:** Started answering, then said that he would rather not, and left 87 A and B blank.
**R9:** Failed to answer parts b and c until prompted, perhaps because answer was zero.
**R10:** No finding.

**Q. 87B - "On how many occasions (if any) have you taken cocaine in any form (including crack)?"**

**R1:** See Q85.
**R2:** No finding.
**R3:** R did not finish reading the question, but rather vigorously said no to any cocaine use.
**R4:** R did not answer for use in "...12 months" and "...30 days" although she reported some use in her lifetime.
**R5:** R responded that she used cocaine on 0 occasions in her lifetime and then skipped the items asking about use in the last 12 months and in the last 30 days.
**R6:** No problem nor hesitation about question--"I don't mess with that stuff." Same comments above about drugs apply to this question as well.
**R7:** No problem with question. Asked R if she thought other Rs would answer drug and alcohol questions truthfully. Said she thought they would for alcohol and pot questions, but not for cocaine.
**R9:** Failed to answer parts b and c until prompted, perhaps because answer was zero.
**R10:** No finding.

**Q. 88 - "Which of the following people live in the same household with you?"**

**R1:** R initially considered her sister in item F and was unsure of classifying other household members. She missed the section "Enter Number". POSSIBLY REORDER CHOICES FROM COMMON CATEGORIES TO THE LESS COMMON. MAYBE COLLAPSE SOME CATEGORIES TO REDUCE THE LIST SOMEWHAT.
**R2:** R was slightly confused by the instruction "ENTER NUMBER WHERE APPLICABLE" but figured it out.
**R3:** R selected c. to include her mother's boyfriend. She was unsure about entering a number for j. to count her sister.
**R4:** No finding.
**R5:** R did not read the instructions aloud. She initially skipped the number of her brothers and sisters she lives with.
**R6:** R initially forgot to enter the numbers for children, siblings and grandparents. Then when she did, she didn't zero fill. She said she didn't enter numbers at first because she didn't see instruction--recommends moving the instruction down closer to the boxes. The zero fill problem--"I just blanked". I asked if an example for filling out boxes should be part of the instruction page in the beginning of the question and she said "yes". R also recommends that the list start with mother and mother types because this applies to most dropouts.
**R7:** R had no problem recording parents and baby. Sisters less clear until she reread and saw category j. Finding category for her sister's baby was more of a challenge. Filled in numbers correctly, but forgot to circle some codes.
**R8:** At start of interview R reported having two cousins in household, but said at this question that he should have reported just one. Selected correct

categories to report two adults (grandmother and aunt).

**R9:** Both a boyfriend and a girlfriend live in household, but R only circled h (unsurprisingly reading this as boyfriend and girlfriend), instead of h and l. COMMENT: THIS QUESTION IS EASILY ANSWERED INCORRECTLY. ONE ALTERNATIVE APPROACH IS TO FIRST ASK ABOUT NUMBER OF PEOPLE IN HOUSEHOLD, AND THEN FOLLOW WITH A QUESTION ABOUT RELATIONSHIP OF EACH PERSON TO RESPONDENT.

**R10:** R wanted time reference. Wanted to know if this Q was referring to "now" or when he was in school. Missed "enter number" column completely, until prompted.

Explanation box at top of page 54: **NOTE: The following two questions pertain to fundamental freedoms of expression. These questions will provide helpful information for the interpretation of survey results. We hope you will answer every question, but you may skip any question you do not wish to answer.**

**R1:** No finding.

**R2:** Not clear to R.

**R3:** R thought that the explanation meant that she had the freedom to answer any way she wanted.

**R4:** R understood the explanation to mean that she has the freedom to express herself and to not have to answer anything she doesn't want to.

**R5:** No finding.

**R6:** R said she doesn't understand why its there at all--"you've already told us questions are voluntary." I reminded her that this questionnaire doesn't include all of the questions and that the statement may make more sense in the real document. Still, she thinks its unnecessary and makes R more scared of questions. When I probed her understanding of the statement, she did question the word "fundamental". She knew freedoms of expression must relate to freedom of speech. "Dropouts or anybody else won't know what you're talking about." When I asked for suggestions to improve statement, she said just to get rid of it.

**R7:** R reread statement several times before asking for clarification on meaning of "fundamental freedoms of expression". R thought it meant free speech. Interpretation was a totally unfamiliar word.

**R8:** Did not understand several terms, including fundamental.

**R9:** Read and appeared to understand.

**R10:** Read it, was unable to define "fundamental freedoms of expression" adequately.

**Q. 99 - "In the past year, about how often have you attended religious services?**

**R1:** R took a while to answer.

**R2:** No finding.

**R3:** No finding.

**R4:** No finding.

**R5:** R did not read the instructions aloud.

**R6:** No hesitation or problems answering question. Feels other R's will answer.

**R7:** When read question asked why the box preceeded it -- religion isn't that private. Said young people don't feel sensitivity of this like older people do. Young people more willing to talk about sex, religion, marital and family stresses.

**R8:** wrote "1 out of year" on the More than once a week line.
**R9:** Answered appropriately.
**R10:** No finding.

Q. 100 - "Do you think of yourself as a religious person?"

**R1:** No finding.
**R2:** R was not sure what was meant by "religious person". She wanted to know if it means someone who goes to church everyday or someone who is very spiritual but does not go to church often.
**R3:** No finding.
**R4:** No finding.
**R5:** R did not read the instructions aloud.
**R6:** Same as above.
**R7:** No findings.
**R8:** Appeared to select correct phrase.
**R9:** Selected No, not at all, yet reported being a Lutheran. No appeared to be chosen because she believed one had to go to church to be a religious person. COMMENT: THIS IDEA MIGHT HAVE BEEN FOSTERED BY PRECEDING QUESTION ON CHURCH ATTENDANCE.
**R10:** No finding.

SUMMARY COMMENTS ON OTHER ISSUES:

R7: General comments: A very cooperative respondent. Wanted to help. She was somewhat intimidated by the terminology in the questionnaire and by having the interview taped. Said it would be nice if the questionnaire could be worded in such a way as to make dropouts feel OK about themselves and their accomplishments. Now, the words and phrases just made her feel stupid. Few opportunities to recognize any good in their lives.

R8: Reading skills: Tutor reported R had reading skills that were about average for the GED class, which had started last September. R read everything quite slowly, misread a number of fairly simple words, and for most of the fairly complex words either asked what they meant, or did not read the word or section.

Questionnaire instructions: R usually read the instructions for a question, although sometimes failed to follow them. He did not consistently follow the circle all that apply or circle one only instructions. In particular, he often seemed prepared to circle more than one item when the instructions were to circle one only. Usually he failed to find more than one item that was appropriate, thus usually giving the appropriate response, although at the cost of some wasted time.

He was also inconsistent in how he gave his answer. For most of the closed questions he wrote his answer on the dotted line by the appropriate number, rather than circling a number.

R missed a number of the skips, and although he sometimes made a comment that suggested he realized a question was inappropriate, he did not go back to look for skip instructions, and would often spend quite a lot of time trying to answer, before leaving it and going onto next question.

R9: Although R indicated that she planned to get a diploma, and hoped to go to the University of Illincis to get this, she was unsure whether she would get a diploma, GED or High School Equivalency from the program.

VII. Language Use: Not relevant: Although R was born in Thailand she had lived in the U.S. since the age of 4, and spoke excellent English.

General comments: Probably the most serious problem with respondent's answering of the questionnaire, is that she was inconsistent when answering questions about the last school she attended. Sometimes she defined this as being the high school she attended prior to dropping out, and at other times the preparatory school she had recently attended for just one month. I suspect that quite a lot of other drop-outs who have attended some type of school since leaving high school will be similarly inconsistent.

More generally, it is worrying that this respondent made numerous quite serious errors, despite being a better reader, more intelligent and more conscientious than the large majority of school drop-outs who will be administered the instrument in the survey. If the other cognitive interviews show a similar pattern, the questionnaire needs a lot of work.

Questionnaire instructions: After answering the first two questions by circling the number, R began circling the appropriate word rather than the number. After I pointed out error she corrected this, although occasionally would revert to circling a word.

Skips: A number of skip instructions were missed. On occasion she would go back and see skip, and then correctly follow it.

Work history: Currently working in a shop repairing, ordering, selling, etc. Had done numerous jobs since leaving high school.

Education: Had not gained GED or High School Equivalency. Had briefly attended a preparatory school since leaving high school, and was hoping to go to a college to get a diploma at some future time.
Reading skills: R was a fairly good reader, although she misread certain words, including fairly simple ones, perhaps due to inattention.

R10: Comments: I think R was high; was obviously in it for the money, but when I was able to bring his attention to the task at hand, he tried to vocalize for me.